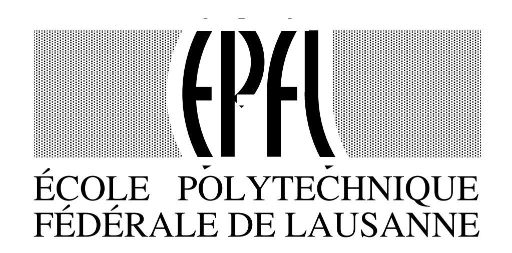
FACULTE ENVIRONNEMENT NATUREL ARCHITECTURAL ET CONSTRUIT (ENAC)

SECTION SCIENCES ET INGENIERIE DE L'ENVIRONNEMENT (SSIE)

LABORATOIRE DES SYSTEMES D'INFORMATION GEOGRAPHIQUE[1] (LaSIG)



ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

TRAVAIL PRATIQUE DE DIPLOME

Conception d'un Système d'Information Géographique (SIG) pour la gestion d'un domaine expérimental.

Application au domaine expérimental viticole du Caudoz, Pully, Suisse.	
Candidat :	Christophe MERTINA
Professeur :	François GOLAY
Responsable :	Régis CALOZ
Encadrement :	Karine PYTHOUD
En collaboration avec la station fédérale de recherche en production végétale (RAC), section viticulture et œnologie, domaine du Caudoz (RVC), Pully.[2]	
	LAUSANNE, février 2004.

[1] LaSIG : EPFL, Section SSIE, Bâtiment GR, CH-1015 Lausanne.

[2] Station fédérale de Changins (RAC), Centre viticole du Caudoz, CH-1009 PULLY.

Table des matières



Que cherchez-vous ?

NUMERO DE CHAPITRE	CE QUE VOUS Y TROUVEREZ
1	Une introduction générale sur le travail et les partenaires
2	Des connaissances sur les SIG et sur le vignoble
3	Une description de la méthode mise en œuvre
4	Une description du fonctionnement de la station
5	L'évaluation des besoins et des attentes des chercheurs
6	Les perspectives de développement du SIG
7	L'analyse de l'existant, les inventaires et les acquisitions
8	La modélisation et la conception du SIG
9	Le prototypage sur logiciel
10	Des exemples d'exploitations, d'utilisations…
11	Quelques considérations sur le traçabilité
12	La conclusion et les perspectives

Contenu du CD

LE CD ANNEXE AU PRESENT RAPPORT CONTIENT TOUT CE QUI A ETE UTILISE POUR REALISER CE TRAVAIL DE DIPLOME, DONT :

Le rapport,	Word
Le prototype SIG, et les fichiers utilisés	Manifold
Les annexes dont le dictionnaire des données,	Word
Un dossier contenant divers fichiers et dossiers utilisés pour le présent travail	Divers
La présentation	Power point
D'autres documents utiles.	Divers logiciels



Résumé

La station fédérale de recherche en production végétale (RAC) section viticulture et oenologie de Caudoz (RVC)[3] réalise des expérimentations à tous niveaux sur la vigne mais aussi sur le vin : elle étudie l'influence de certains paramètres déterminants sur le vignoble. Une multitude d'essais, de tests, et d'expériences sont menés chaque année et suivis pour une partie d'entre elles dans le temps. Des conclusions peuvent ainsi être apportées sur la base d'analyses statistiques issues de traitements des données.

Il apparaît clairement que **les données** sont nombreuses, diversifiées mais aussi très hétérogènes, non centralisées, et surtout non référencées au territoire d'où pourtant elles proviennent pour la majorité d'entre elles. Il en découle une démarche de travail adapté à un fonctionnement complexe structuré à plusieurs niveaux au sein d'une même institution. La mise en place d'un outil SIG a déjà été envisagée mais jamais réalisée.

Ce travail vise donc à **concevoir un SIG qui permette la gestion du domaine expérimental**. La problématique est donc spécifique à un thème précis, celui de la viticulture à la RVC, mais elle est facilement extrapolable à d'autres domaines d'activités semblables, c'est-à-dire fonctionnant selon la même structuration de données.

Ce rapport expose les démarches d'analyse, de conception et de réalisation du prototype de SIG. Il aborde aussi les problèmes, avantages et inconvénients d'un tel système, et il met en évidence les possibilités et problématiques d'extension. Il donne quelques recommandations et aspects importants liés à la problématique. Pour terminer, la notion de traçabilité, très à la mode en ces temps de « suivi alimentaire », est apparentée au sujet.

[3] Afin de ne pas confondre les lieux et les spécialisations au sein de la RAC, l'abréviation RVC est employée pour spécifier le domaine étudié, à savoir le domaine du Caudoz à Pully, spécialisé dans la viticulture et l'œnologie.

Remerciements

Je tiens à remercier toutes les personnes qui m'ont permis de mener à bien ce travail.

En particulier :

M. François Golay
professeur au LaSIG
directeur du travail de diplôme

M. Régis Caloz
collaborateur scientifique au LaSIG
responsable du travail de diplôme

Mme Karine Pythoud
collaboratrice au LaSIG
encadrement du travail de diplôme

M. Vivian Zufferey
ingénieur agronome au service viticulture–œnologie de la RAC
collaborateur externe en charge du projet

M. François Murisier
directeur du service viticulture–œnologie de la RAC
directeur externe en charge du projet

L'ensemble des collaborateurs du LaSIG
Pour leur aide aux niveaux « informatique » et « modélisation »

L'ensemble des collaborateurs de la RVC
Pour leur contribution à l'acquisition des données.



Liste des abréviations

ABREVIATION	DESIGNATION
DCA	Diagramme Conceptuel d'Activité
DFD	Diagramme de Flux de Données
Id	Indice de Dégustation
ID	Identifiant
Iv	Indice de Vinification
LaSIG	Laboratoire des Systèmes d'Informations Géographiques, EPFL.
MCD	Modèle Conceptuel de Données
MLD	Modèle Logique de Données
RAC	station de recherches en production végétale de Changins / Nyons
RHM	Réserve Hydrique Modale
RU	Réserve Utile
RVC	centre de Recherche Viticole du Caudoz
SIG	Système d'Informations Géographique
SIRS	Système d'Information à Référence Spatiale (synonyme de SIG)
TPD	Travail Pratique de Diplôme

Liste des figures

Liste des tableaux

Liste des annexes

NUMERO	TITRE
1	Dictionnaire des données
2	Manuel d'utilisation simplifié du SIG
3	Principe de lecture des codes des unités de sol
4	Structure générale du Modèle Conceptuel de Données (MCD)
5	Modèle conceptuel des données
6	Modèle Logique des Données (MLD)
7	Hiérarchisation du projet sous MANIFOLD
8	Les tables OBJETS, BASES et ESSAIS
9	Quelques illustrations des résultats de la digitalisation.



Dans ce chapitre :
Description sommaire, localisation, missions du mandataire : la RVC.
Les objectifs du travail, et la problématique associée.
Pourquoi un tel travail ?
Les étapes du déroulement du travail.

1 Introduction

1.1 Qu'est ce que la station fédérale de recherche viticole de Caudoz (RVC) ?[4]

Située sur la commune de Pully (Cf. fig. 1 et 2 : localisation de la RVC), **la RVC** réalise des recherches sur le vignoble. Ces travaux sont en étroite collaboration avec *la station fédérale de recherche en production végétale* de Changins (RAC), près de Nyon, dont elle est une section parmi tant d'autres. **Sa mission** (Cf. table. 1 : les missions de la RAC) est d'étudier la vigne au sens large du terme. Une multitude de données sont ainsi collectées, elles sont de natures diversifiées du point de vue de leur source et de leur type mais aussi très nombreuses. Quatre collaborateurs sont postés à Caudoz pour ce qui est de la partie *viticulture et oenologie*, la majorité des collaborateurs étant à la RAC. Leurs fonctions sont essentiellement liées à la recherche en partenariat avec différentes unités nationales et internationales : ce sont donc des chercheurs. Ils travaillent en suivant dans le temps des **essais** réalisés sur des parcelles du **domaine** de 4 hectares (Cf. fig. 3: plan du domaine expérimental de Caudoz). Ce dernier comprend également des bureaux, d'une station météo d'importance nationale et un petit laboratoire.
Il a déjà été envisagé la mise en place d'un système centralisé, fédérateur, permettant de traiter les données selon les besoins des chercheurs. Pour ce système, l'évocation d'un SIG n'est pas nouvelle. Depuis ces réflexions, le choix d'une base de données traditionnelle a été réalisé et a fait ses preuves pour un secteur d'activité, les autres secteurs d'activité conservant les bonnes vielles méthodes des cahiers de données indépendants mais si fiables ! Aujourd'hui, l'opportunité de la réalisation du projet cantonal *terroirs viticoles vaudois* a permis de se pencher un peu plus sur cette problématique :

- Est-il utile de mettre en place un SIG à la RVC ?
- Comment le rendre efficace, adapté, fonctionnel sans complexité, en restant proche d'un comportement connu ?
- Qu'apporte le SIG, quels sont les avantages et les inconvénients d'un tel système ?

[4] Site Internet : www.admin.ch/sar/rac, section viticulture et œnologie.

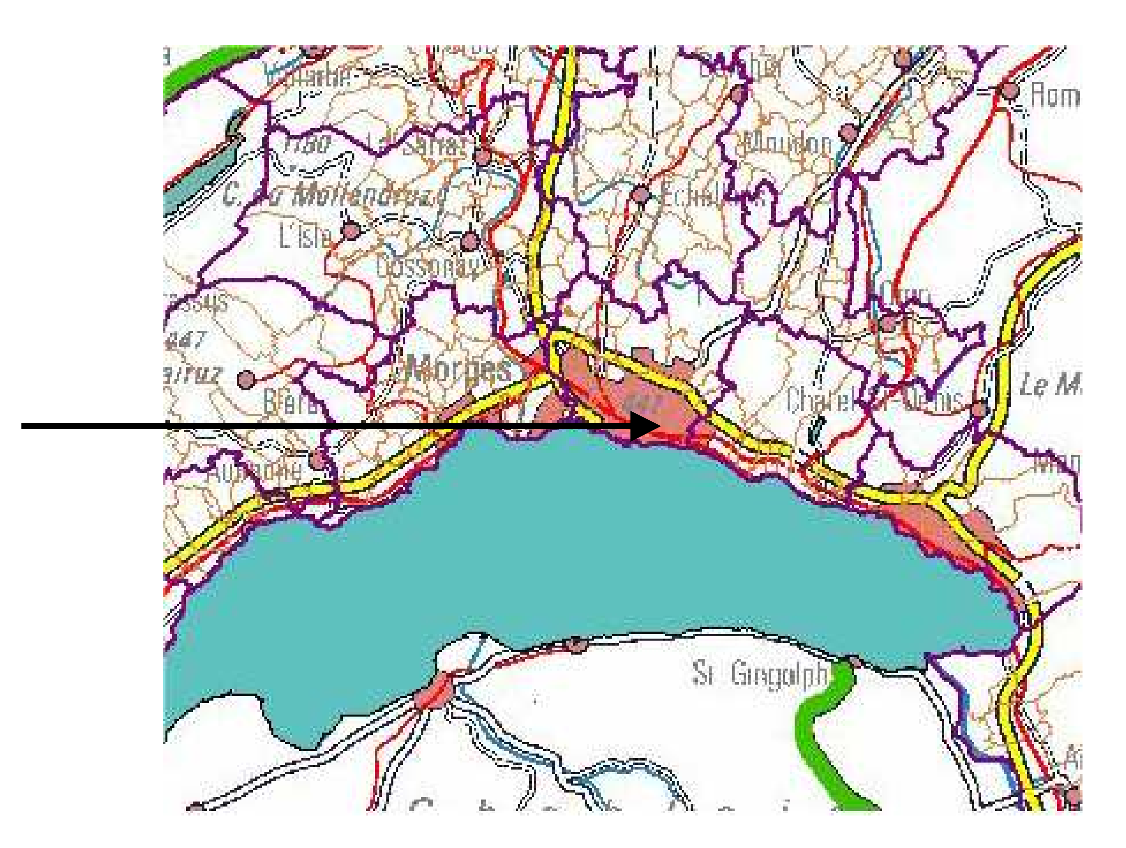


Figure 1 : situation approximative de la RVC, au bord du lac Léman, VD, Suisse.

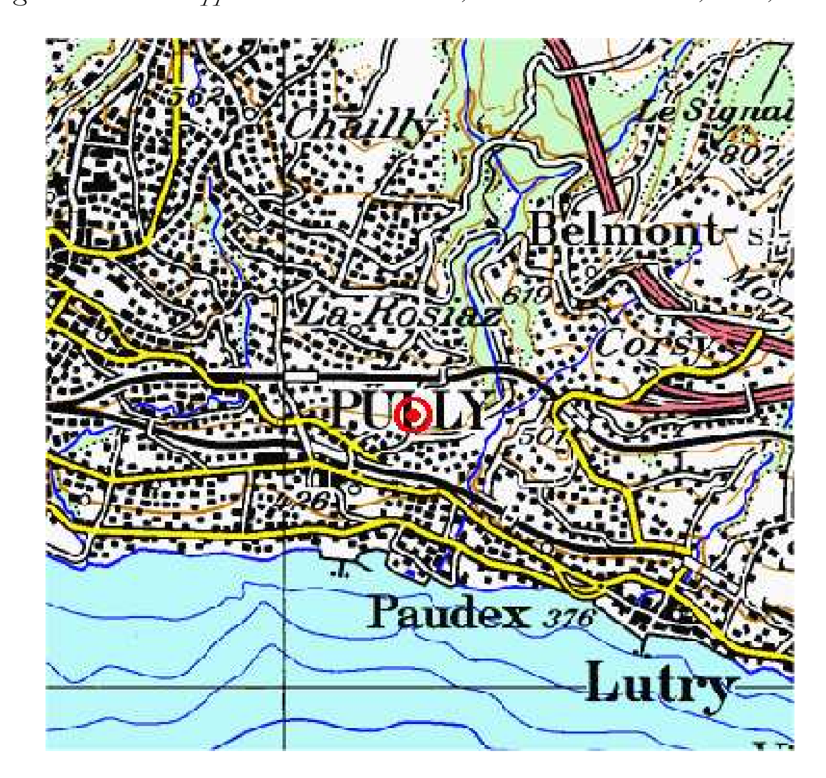


Figure 2 : situation de la RVC au niveau de la commune de Pully (point de localisation).



Figure 3 : plan des parcelles expérimentales (avec leur numéro) du domaine de la RVC avec les locaux.

Missions

1 : Sélection et étude des cépages

Sélection clonale des principales variétés cultivées en Suisse (Chasselas, Pinot et Gamay) et des « spécialités » valaisannes dans le but de repérer les meilleurs clones, de les diffuser (certification) et de conserver une diversité génétique.

Création et sélection de nouveaux cépages (métis de vinifera et hybrides interspécifiques), dans le but d'accroître la palette des cépages, très limitée dans certaines situations climatiques et d'obtenir un bon niveau de résistance aux maladies fongiques.

2 : Adaptation des cépages et porte-greffe dans différentes situations pédoclimatiques suisse. Techniques culturales de la vigne

Optimalisation des modes de conduite de la vigne en vue d'assurer une production maîtrisée et une qualité optimale et de réduire les coûts de production. Gestion de l'entretien du sol et de la fumure avec pour objectif de proposer des techniques permettant de gérer le stress hydro-azoté dû à l'enherbement et de garantir la qualité des vins. 3 : Technologie, microbiologie et chimie des vins Mise au point de techniques de vinification avec pour objectif de diversifier les produits viticoles et d'améliorer la qualité des vins. Testage et sélection de levures et de bactéries en vue d'améliorer les fermentations et la qualité des produits. Contrôle analytique des moûts et des vins, notamment en relation avec l'exportation des vins suisses. Source : Site Internet de la RAC.

Table 1 : Les missions de la RVC : trois missions essentielles déclinées en études.

1.2 Objectif du travail pratique de diplôme

L'objectif de ce travail de diplôme est de concevoir un système d'information géographique (SIG) permettant la gestion de la RVC. Il s'agit dans une première étape, de proposer un modèle de base de données à référence spatiale susceptible d'être exploité pour la gestion du domaine de recherche, c'est à dire se devant d'être adapté au fonctionnement de la station, et dans une deuxième étape, de réaliser un prototype pour le stockage structuré des informations collectées mais aussi pour leur utilisation appropriée.

Une attention particulière doit être apportée aux points suivants :

- Cerner les attentes des chercheurs en matière de gestion de données
- Proposer des applications potentielles de la base de données géographique au domaine
- Etendre la réflexion à la notion de la traçabilité des produits.

1.3 Problématique

La **problématique** de mise en œuvre d'un SIG dans une station de recherche telle que la RVC se résume aux points particuliers suivants que l'on retrouve dans bons nombres d'organisations :

- Multitude de données
- Grandes variétés de données très diversifiées
- Multiple gestion et utilisation des données,
- Mobilité des collaborateurs,
- Plusieurs sites de prospection,
- Projets / études temporaires et permanents,
- Fonctionnement actuel ayant fait ses preuves,
- Introduction d'un SIG en collaboration avec d'autres bases de données,
- Besoins et attentes des chercheurs pas bien définis au regard d'un SIG,
- Etc…

1.4 Pourquoi un tel travail pratique de diplôme ?

Le besoin de lier les informations au terrain, auquel elles se rapportent pour une grande partie d'entres-elles, devient actuellement à la mode pour des raisons simples : convivialité des systèmes SIG, meilleures structurations des données, visualisation spatiale, possibilités multiples de traitements et de fonctions. Fini le temps des listing difficilement agréables !. **Les attentes des chercheurs** vont dans ce sens : il leur paraît indispensable d'effectuer leurs travaux en relation avec la dimension spatiale et ceci pour plusieurs raisons (allant de la simple visualisation des parcelles d'essais jusqu'au besoin d'établir des cartes thématiques ou des analyses spatiales). Les données à leur disposition sont nombreuses et diversifiées ; elles sont intégrées dans un schéma de fonctionnement oú le chercheur choisit les données dont il a besoin pour mener à bien ses études et projets. On peut appeler ceci un *fonctionnement à la carte*. Les données ont besoin en outre d'être centralisées. **Les perspectives de développement futur** sont nombreuses : du simple système d'information géographique (SIG) au système global de gestion intégrale des données avec possibilités de traitements et d'analyses allant vers une mise en valeur des données. **Le SIG projeté** devrait comporter de nombreux avantages en plus du système actuel de fonctionnement, mais apporter aussi ses limites et ses contraintes d'exploitation et d'utilisation.
Le résultat du travail sera donc un prototype de SIG qui permettra de répondre au plus grand nombre des besoins des chercheurs, et de mettre en évidence les avantages et problèmes découlant de l'utilisation d'un SIG pour une telle organisation de recherche.
Tout l'intérêt du travail réside en cette interrogation :

Un SIG pour Caudoz? Pourquoi ? Comment ?

1.5 Déroulement du travail

Les étapes de travail sont les suivantes :

- Interview des chercheurs en charges du domaine,
- Inventaire des données et des besoins,
- Modélisation / Conception d'un prototype SIG à but multiple,
- Essai « grandeur nature », conclusion, propositions.

Résumé des points importants :

La RVC (le mandataire) réalise des essais vitivinicoles sur un domaine de recherche.

L'objectif du travail de diplôme est de s'interroger sur l'utilité, la mise en œuvre, la conception et le prototypage d'un SIG pour la gestion du domaine.

La problématique est complexe mais « traditionnelle » pour ce genre d'entreprise.



Dans ce chapitre :

Tous les éléments nécessaires de connaître pour ce travail concernant les SIG et le VIN.

2 Quelques généralités

2.1 Sur les SIG

2.1.1 Définition de SIG

Simplement un SIG est un système informatique permettant **de « travailler » avec des données référencées** au territoire, pour la majorité d'entre elles.

Plus exactement un SIG (système d'information géographique) ou système d'information à référence spatiale (SIRS), est un ensemble d'équipements informatiques, de logiciels et de méthodes. Il rend possible la saisie, la validation, le stockage et l'exploitation des données dont la majorité est spatialement référencée. Les fonctions associées permettent la simulation du comportement d'un phénomène naturel, le traitement d'informations pour la **gestion** d'une entreprise et l'aide à la décision. Un SIG fonctionne souvent comme **une mémoire des informations liées au territoire.** *[CALOZ, 1994]*.

Le principe d'un SIG est de mettre en relation ces informations (par le biais d'attributs) afin d'identifier, de structurer, de simuler et de cartographier des résultats pour visualiser, comprendre et aider.

2.1.2 Intérêts principaux d'un SIG

- Faciliter les études et les prises de décision,
- Editer des cartes, des documents …
- Connaître précisément l'état des lieux de l'espace géographique pour éclairer en temps réel les réflexions,
- Croiser des informations géographiques, de tous types dans un référentiel géographique commun pour permettre des analyses et visualiser les phénomènes,
- Permettre la superposition de cartographies différentes offrant la possibilité de rapprochements de renseignements,
- Modéliser des phénomènes pour effectuer des simulations,
- Doter les collaborateurs d'un outil moderne pour la gestion des territoires,

2.1.3 Quelques notions importantes liées aux SIG

Le tableau 2 ci-dessous présente succinctement **les différents éléments que l'on rencontre** rapidement lorsque l'on est confronté aux SIG. *[GOLAY, 1999* & *2000]*.

NOTION	EXPLICATION
Attribut	C'est une information rattachée à un élément
Base de données, relationnelle	Une base de données est constituée de l'ensemble des informations nécessaires au fonctionnement d'une entreprise ou d'une organisation, ensemble dont la gestion est assurée par un logiciel appelé système de gestion de base de données (SGBD). Moteur de la base de données, le SGBD permet de décrire les données et de les travailler, de maintenir leur intégrité, de gérer les transferts, d'assurer la sécurité et le contrôle des accès. Dans le cadre des SIG et dans la majorité des cas, le SGBD associé est de type relationnel, il est intégré directement au SIG. Une base de données est dite relationnelle parce que son fonctionnement est basé sur des règles mathématiques rigoureuses dénommées algèbre relationnel
Carte thématique	Sont des cartes réunissant des données par sujets identiques ou par types de résultats : le thème de la carte.
Couche	Une couche représente une catégorie d'informations appartenant à la même thématique. Exemple : couche de la couverture du sol, couche des bâtiments…
DFD	Le diagramme de flux de données, montre les relations, les chemins, les traitements… autrement dit le parcours des données au sein d'une entreprise.
Dictionnaire	Le partage d'informations entre de nombreux partenaires exige une approche globale et concertée, ce qui implique une référence spatiale commune et surtout des définitions partagées.
Digitalisation	Comme l'on passe d'une figure papier à une figure numérique « raster » par le scanner, on passe de l'image « raster » à la vectorielle par la digitalisation.
Donnée	C'est une information, pouvant aussi être spatiale c'est à dire « cartographique ».
Donnée référencée	C'est une donnée liée au territoire, elle dispose donc de coordonnées.
Donnée référentielle	C'est une donnée à laquelle d'autres intéressés que le gérant rattachent des informations
Echelle	Elle indique le degré de détail de représentation de l'information spatiale
MCD	MCD signifie Modèle Conceptuel de Données : décrit les composantes (les objets et leurs liaisons) de la réalité qui seront traitées par le système, indépendamment des aspects logiciels et matériels. Ce modèle sert de fondement à la conception d'une base de données
Méta donnée	Une méta donnée désigne l'ensemble des informations venant se greffer sur les données (on dit couramment : des données sur les données)
Mise à jour	On dit souvent : « Tenir à jour » les données, c'est à dire que à intervalle de temps réguliers, la base de données se doit d'être en accord avec la réalité
Modèle	Il représente une simplification d'une réalité complexe. Ainsi, une information décrit un modèle de la réalité
Normalisation	Les données doivent être normées afin qu'elles soient cohérentes, échangeables, en un mot en harmonie.
Saisie de données	Représente le fait d'acquérir des données, attributs et données référencées. Il s'agit dans cette étape de « remplir » la base de données.
SGBD	Un système de gestion de base de données (SGBD), intégré au SIG, est un ensemble de logiciels permettant le stockage, la gestion et la récupération des données contenues dans une base de données. Le SGBD est le moteur du SIG.
Table	C'est un listing contenant des informations rattachées à un élément
Technologie raster	L'espace terrain est divisé en pixels auxquels on associe une valeur. A

	chaque valeur on associe une valeur de gris de couleur, ce qui donne une image : le territoire.
Technologie vecteur	Le territoire est représenté par des primitives géographiques.(a contrario de la technologie raster)
Traitement	Il regroupe toutes les opérations nécessaires pour obtenir une information pertinente à partir des données de la base

Table 2 : les notions importantes des SIG et leurs explications. .

2.1.4 Les enjeux du marché actuellement

Comme tous les enjeux de marché, ils résultent du rapport qualité / prix.
La qualité des différents logiciels doit être appréhendée en fonction des besoins des utilisateurs. La concurrence est rude. Les prix on tendance à baisser pour une qualité et des fonctionnalités toujours plus perfectionnées.
Les possibilités de mise en place existent, sont nombreuses, rapides et réellement intéressantes du point de vue de leur service rendu et de leur fonctionnalité. Nous tendons à une généralisation d'ici à quelques années des systèmes SIG. Ils vont remplacer peu à peu les « données papier » et s'intégrer dans notre quotidien, tout comme l'a remarquablement fait l'Internet. Une multitude d'entreprises proposent leurs services pour ce genre de travaux.

2.1.5 Les avantages et les inconvénients des SIG

2.1.5.1 Les avantages

L'apport de la mise en place d'un système SIG est non négligeable Ci dessous sont dressés les points les plus importants qui suffisent **à prouver qu'un SIG peut avoir sa place dans une organisation de recherche** : *[MERTINA,2003]*.

Accès à la dimension spatiale :

- Visualisation du territoire,
- Possibilité d'édition de plans et de cartes,
- Possibilité d'édition de cartes thématiques[5],
- Possibilité de discuter sur des problématiques à référence spatiale, carte en main.
- Limitation des sources d'erreurs par rapport à un système dit codé, du fait de la visualisation, discussion, édition de documents.
- Création de synthèses et de documents de synthèse (pré préparés ou non) contenant d'une part les informations tablées et d'autre part leur représentation spatiale ; (titre, tables, légendes, coordonnées, zoom, orientation…).

Conservation du fonctionnement actuel :

- organisation générale maintenue,
- conservation des données existantes,
- continuité du mode de diffusion des données (publication, synthèses, cahiers…),

[5] Cartes réunissant des données par sujets identiques ou par types de résultats.

- habitudes des collaborateurs conservées sans changement fondamental ,

Réponse aux besoins et aux attentes de la plupart des collaborateurs:

- consultation possible des données en vue d'informations,
- requêtes, interrogations possibles,
- gestion des données de manière relationnelle[6]
- historique conservé,
- possibilité de traitements des données (statistiques, corrélations …),
- possibilité de mise en valeur des données (cartes thématiques…),
- connexion éventuelle du SIG à d'autres bases de données externes,

Réponse contre l'individualisme,

- données de base communes à l'ensemble des collaborateurs,
- pas de problème de sources de données différentes,

Perspectives de développement particulièrement intéressantes et nombreuses :

- nouveautés dans la gestion des données, dans leur mise en valeur et leur traitement,
- apports nouveaux jusque là ignorés,
- meilleur rendement du fait de la gestion commune des données de base,
- réponses aux lacunes du système actuel de fonctionnement.

D'autres avantages existent mais sont plus spécifiques à des circonstances précises.

2.1.5.2 Inconvénients

- Nécessité d'une motivation et d'une formation des collaborateurs au nouveau système,
- Manifestation d'une volonté de changement en vue d'une amélioration,
- Formation d'une personne pour la maintenance du SIG ou mise en place d'un mandat de gestion
- Nécessité de travailler selon une méthode définie, en concertation, afin d'obtenir des données communes utilisables, fiables, à jour…
- Léger changement des habitudes de travail,
- Nécessité d'une autonomie des collaborateurs face au SIG, pour promouvoir un travail efficace.

[MERTINA,2003].
On remarque que les inconvénients sont plutôt d'ordre **humain**, avant d'être opérationnels ou techniques. Des solutions à ces quelques inconvénients peuvent être apportées par un **accompagnement au changement**, une explication du « pourquoi changer ? », et de la motivation.

[6] Une base de données est dite relationnelle lorsque son fonctionnement est basé sur des règles mathématiques rigoureuses dénommées algèbre relationnel.

2.2 Sur le vignoble

« Le vin est infini dans sa diversité. Il n'existe pas deux années, deux caves, deux endroits produisant un vin identique et c'est la nature du vin que d'exprimer plus ou moins clairement les conditions de sa production et l'état de ses matières premières. ». *[DUBOIS, 1996]*. Cette partie s'inspire principalement de cet ouvrage ainsi que du site Internet *[De la vigne au vin: http://users.swing.be/john/vignevin/index.htm]* et du guide du vin *[DUMAY, 1967]*.

2.2.1 Un peu d'histoire

Bien des légendes évoquent les origines du vin. On sait toutefois que la plus ancienne viticulture se situe en Transcaucasie (entre la mer Caspienne et la mer Noire), et que les études ont montré scientifiquement que des vases vinaires dataient de **1500 av. J.C**. A l'origine, le vignoble n'existe pas. Il résulte de l'art de la culture de la vigne sauvage *vitis vinifera.* Cette culture de la vigne, que l'on nomme aussi **conduite de la vigne**, cet art est décrit au fil des années, des siècles par les Grecs, les Romains, les Egyptiens, les Hébreux. C'est avec les colonisations grecques (-500), puis romaine (-200) puis par la suite l'avènement de la foie chrétienne (suite à la naissance de Jésus Christ) que la culture de la vigne va se répandre à travers l'Europe, puis le monde. Le rite chrétien de l'eucharistie renforce cette culture, les monastères s'entourent d'enclos réservés aux vignobles. L'art de culture, de transformation, de vinification se poursuit, s'étoffe et se diversifie. C'est ainsi que peu après Jésus-Christ, un moine, Dom Pierre Pérignon, applique les procédés de la champagnisation, ce qui constitua, sans le savoir, l'avènement des fermentations et des vins pétillants, dont le champagne est le plus connu. Cependant, dans les pays où le vin ne se consomme pas, pays arabes, Moyen-Orient, où la loi coranique s'exerce, la culture de la vigne s'étend malgré tout du fait de la simple consommation des raisins et de son jus. Le monde entier se voit ensuite imposer la culture de la vigne et ses modes de conduite suite aux différentes colonisations. Ceci favorise alors une fois de plus **des variétés, des identités uniques** qui s'apprécient encore aujourd'hui : c'est tout ce qui fait la particularité du vin : **son unicité**.

2.2.2 Et maintenant ?

Et maintenant ? Et bien les temps ont changé. Les critères de productivité, simplicité, uniformité, rendement, tentent de l'emporter au profit d'une guerre des prix, et des **marchés** qui s'élargissent par la venue de vins étrangers appréciés.. A contrario, le « business » du vin doit répondre à des fins de qualité, d'image, d'identité : il ne doit pas considérer l'uniformité des vins. C'est pour cette raison que florissent les labels, les appellations d'origines contrôlées, **des terroirs** …. Un véritable combat a lieu. Un combat entre les procédés d'uniformisation du monde et les particularités locales qui sont à la clé du vin, sans lesquelles le vin ressemblerait à de l'eau.
En vue de garantir cette **panoplie diversifiée** et d'en créer d'autres, d'améliorer l'existant, de combattre les maladies, d'améliorer les résistances, la qualité, la productivité, les critères gustatifs, hygiéniques, nutritionnels, etc., **des organismes de recherche en production viticole** ont vu le jour il y a plus d'un siècle. Leur expérimentation vise à garantir et comprendre les acquis ancestraux, gérer les menaces actuelles et prévoir l'avenir d'un produit presque vieux comme l'*Homme* civilisé.

2.2.3 Quelques connaissances utiles

2.2.3.1 Terminologie

Voici, à la table 3 ci-dessous, quelques définitions de termes utiles à la compréhension du **monde du vin**, de la vigne, de la viticulture, du vignoble…*[DUBOIS, 1996]*.

TERME	DEFINITION
Cep	Pied de vigne

Cépage	Plant de vigne cultivé d'une certaine manière, il en existe plus de cinq mille variétés dont seulement cinquante intéressent l'amateur de vin. Pour ce qui est des variétés européennes, elles descendent toutes du même ancêtre : vitis vinifera. Généralement le plan de vigne doit être considéré dans sa spécificité.
Clone	Ensemble, lignée d'individus issus de la multiplication végétative d'un individu initial
Coteau	Versant d'un plateau, d'une colline
Encépagement	Peuplement d'un vignoble
Greffage	Procédé employé pour reproduire la vigne qui consiste à fixer un greffon (fragment d'un cépage que l'on désire reproduire) sur un porte greffe
Greffe	Restauration d'un organe par implantation d'une portion de tissu empruntée à une autre partie du corps ou à un autre sujet
Phylloxéra	Puceron minuscule parasite des racines de la vigne et son ennemi le plus ravageur. Par extension, maladie de la vigne due à cet insecte
Porte-greffe	Sujet sur lequel on fixe des greffons
Sarment	Bois de la vigne, tige ligneuse et grimpante poussant chaque année
Vitis vinifera	Variété de vigne sauvage, porteuse de raisin pouvant donner du vin

Table 3 : définitions des principaux termes viticoles

2.2.3.2 De la vigne au vin

La culture de la vigne est une **œuvre** de patience. Il, faut en effet attendre plusieurs années avant de pouvoir obtenir un vin de qualité. Choix de cépage, encépagement, clonage, greffage, entretien, taille, fumure… ponctuent la vie de la vigne, puis les vendanges. la vinification, la mise en cave, le suivi, la réflexion…celle du vigneron. Passer de la vigne au vin résulte d'un art. Ainsi on parle de la **création d'un vignoble** et non de sa plantation. Les étapes sont nombreuses. On les regroupe en trois parties principales :

- la culture de la vigne : de la nature aux baies,
- les vendanges et la vinification : des baies au jus,
- le suivi de l'élaboration du vin : du jus au vin.

2.2.3.3 Le vignoble

Le vignoble désigne un lieu, précis, une région tout au plus. Il est le résultat de **l'unicité de la vigne** et par extension du vin. Il dépend :

- du cycle végétatif de la vigne elle-même
- de la nature et exposition du terrain,
- de la façon dont la vigne est conduite.

C'est ainsi que sont nés les **grands vins**, Français pour la plupart (la France produit a elle seule le quart de la production mondiale en vin ce qui en fait le premier producteur mondial) connus du grand public, dont on peut citer :

- Côte du Rhône
- Bordeaux
- Bourgogne
- Beaujolais
- Champagne
- Vin du Jura
- Vin d'Alsace

Ces vins portent le nom de leur **région** respective : ils y sont apparentés.

2.2.3.4 L'encépagement et les principaux cépages

L'**encépagement** est caractérisé par le cépage et par la densité de plantation. N'importe quel cépage ne peut convenir à n'importe quel terrain. Aussi est–il choisi en fonction non seulement de la nature du vin à produire, mais aussi des **paramètres locaux**, qui constitueront par la suite un vignoble. Les cépages sont greffés pour la grande majorité d'entres eux. (Cf. § 2.2.3.5 : Les greffes)
La table 4 ci-dessous énumère les quelques **cépages incontournables** : *[DUBOIS, 1996]* :

CEPAGE	CARACTERE
Chardonnay	Chasselas à peau fine
Chasselas	Cépage réputé, originaire de la Saône et Loire, donnant des raisins blancs de table
Fendant	Vin blanc Suisse du Valais
Gamay	Cépage noir donnant des vins rouges
Gewurztraminer	Cépage blanc donnant des vins parfumés, surtout cultivé dans l'Est de la France
Merlot	Cépage, le plus souvent rouge, cultivé dans le Bordelais
Muscat	Cépage dont les baies ont une saveur musquée caractéristique
Pinot noir	Cépage français renommé, cultivé notamment en Bourgogne
Riesling	Cépage blanc donnant de petits raisins, cultivé sur les bords du Rhin et dans d'autres régions du monde
Sylvaner	Cépage blanc cultivé dans l'Est de la France, l'Autriche, la Suisse et Allemagne

Table 4 : les principaux cépages incontournables.

2.2.3.5 Les greffes

L'utilité de la greffe réside dans le fait que la vigne peut ou plutôt doit **être améliorée** afin de résister notamment aux maladies et de favoriser un enracinement robuste. Les complications commencent sérieusement : c'est un professionnel qui greffe, ceci est un art qui demande réflexion et application. Un exemple bien connu des vignerons est celui du phylloxéra[7]. Comme les racines de la vigne européenne ne supportent pas les piqûres du phylloxéra, elles doivent être « remplacées » grâce au greffage par des racines américaines.

7 Phylloxéra : Puceron minuscule parasite des racines de la vigne et son ennemi le plus ravageur. Par extension, maladie de la vigne due à cet insecte.

2.2.3.6 La conduite de la vigne

La **conduite** de la vigne représente tout un **art** qui se choisi au départ avec l'encépagement et qui se poursuit au fil du temps avec l'**entretien**.
La conduite est encore une fois un choix qui se réalise par le vigneron en fonction de critères le plus souvent issus d'une connaissance du cépage utilisé, de la terre fournie et du travail accumulé.

Nous citerons les **interventions** suivantes qui permettent la conduite de la vigne :

- La culture
- La fertilisation / les traitements
- La taille

Changer de mode de conduite, c'est changer un ensemble d'éléments qui constituent, pris dans leur ensemble, le vin. Aussi dit-on que chaque cep à sa propre individualité. Il est important d'adapter les techniques culturales, notamment le mode de conduite, au cépage et à la région afin de pouvoir élaborer **un vin de caractère**.

2.2.3.7 Les éléments ayant un effet sur la vigne et par extension sur son mode de conduite

Ces éléments sont nombreux. Si l'on laisse de côté le mode de conduite résultant du choix du vigneron, il est des paramètres que le vigneron doit apprendre à connaître pour faire en sorte que son vin puise tout son caractère dans des **particularités** rarement reproduites.

Le premier paramètre est **le temps** : il conduit à l'élaboration de différents vins dont les meilleurs sont appelés **millésimes**. Il n'existe pas deux années identiques qui donneront le même vin.

Le deuxième paramètre est **l'espace** : caractérisé par des conditions de localisation qui lui confère aussi bien des facteurs climatiques que liés au relief. On citera par exemple :

- L'intensité lumineuse
- La température
- Les précipitations
- La pente
- L'altitude
- L'environnement : montagneux, plat, proche d'un lac…
- Le macroclimat
- Les microclimats
- …

Tous ces paramètres ont alors une influence sur la **composition chimique** du vin, influence qui manifestera ces effets lors de chaque étape qui mènera le vin à son unicité. Ce sont principalement :

- La teneur en sucre
- L'acidité
- L'encavage
- La mise en bouteille

- …

2.2.3.8 Q'est qui fait la typicité d'un vin ?

La typicité d'un vin intervient lorsqu'un groupe de vins se ressemblent. On distingue alors de grandes **familles** qui sont souvent liées à leur région de provenance, Cf. § 2.2.3.3 : Le vignoble.
On remarque que le **climat** intervient de manière déterminante dans **l'expression des caractères** et la typicité des vins. La table 5 ci-dessous donne quelques informations sur la typicité des vins en fonction de leur grande région d'appartenance. *[SITE Internet : De la vigne au vin]*.

Dans les régions chaudes et ensoleillées, la maturité sera élevée, les raisins seront alors riches en sucres et peu acides. On obtiendra alors des vins chaleureux et corsés dont la typicité s'exprimera le mieux dans les vins rouges. Dans les régions plus froides et moins ensoleillées, la maturité sera plus faible, les raisins seront alors moins riches en sucres et plus acides, ils donneront des vins fermes, solides et vifs dont les rouges peuvent être assez tanniques. Dans les régions encore plus froides, à la limite zone de culture de la vigne, la production sera essentiellement une production de vins blancs secs et nerveux. Au-delà de cette limite climatique, le vin obtenu sera très acide et léger en alcool, il ne peut alors qu'être exceptionnellement consommé en l'état et doit être transformé, c'est le cas par exemple des productions réputées de Champagne.

Table 5 : la typicité des vins selon leur région d'appartenance.

A partir de ces généralités, la réalité peut être différente car les variations climatiques seront plus complexes et feront intervenir les climats locaux et les microclimats. Ainsi, c'est la résultante de tous ces paramètres liés au climat qui donnera finalement le caractère particulier des vins issus d'un terroir donné.

A ce terme de cette description, nous pouvons dire que nous avons juste effleuré le domaine de la viticulture. Mais notre travail consiste à la mise en œuvre d'un SIG et non d'un mode de conduite pour la vigne. **Le SIG qui devra tenir compte des paramètres évoqués dans ce chapitre**.

Résumé des points importants :

Un SIG est un système permettant de *travailler* avec des données référencées au territoire.

On parle du caractère, de la personnalité, de l'unicité d'un vin comme celui d'un être.

Conduire un vignoble est un art aboutissant à une œuvre.

Le SIG devra intégrer ces notions afin d'être adapté à la situation.



Dans ce chapitre :

L'explication de la méthode choisie et mise en œuvre pour le travail
Le rôle du prototype.
Discussion des aspects organisationnels, techniques et humains pour le projet.

3 Méthodologie

3.1 Mise en œuvre

Afin de mener à bien un projet, il est indispensable de mettre en place **une méthode**. Cette méthode doit nous permettre d'élaborer un SIG fonctionnel, qui réponde aux besoins et attentes des chercheurs et qui soit résistant dans le temps.

Généralement, la méthode d'élaboration du SIG se base sur les principes suivants : *[GOLAY, 1999* & *2000]*

- Délimitation du cadre de l'étude
- Démarche allant du général au particulier, du plus concret au plus abstrait,
- Démarche par étapes,
- Réflexion indépendante de toute solution informatique,
- Prise en compte de la réalité existante.

C'est selon cette démarche que ce travail de diplôme progresse : elle représente la **démarche conventionnelle traditionnellement** utilisée pour la mise en œuvre de SIG. *[GOLAY, 1999* & *2000]*
Les étapes sont les suivantes, que l'on retrouve en tète de chapitre :

1. Analyse du fonctionnement actuel de l'entreprise: description
2. Analyse de l'existant et inventaire des données : recherche
3. Evaluation des besoins et des attentes : discussion
4. Evaluation des perspectives de développement du SIG : proposition
5. Conception du SIG : modélisation
6. Réalisation du SIG : implantation
7. Illustration du fonctionnement, exploitations, extensions: utilisation

3.2 Rôle du prototype

Un prototype est un **système exploratoire**. Ainsi il n'est pas définitif, c'est à dire qu'il n'est le plus souvent pas retenu comme tel dans l'élaboration du système définitif mais s'en inspire fortement. Le prototype doit surtout montrer les problématiques, les besoins, les enjeux, mais aussi servir de formation et de premier contact avec les collaborateurs. Il se doit d'être évolutif pour permettre son amélioration continue en vue d'une optimalisation. Le présent travail s'implique entièrement dans cette perspective.

3.3 Prise en compte des facteurs techniques, organisationnels et humains

L'ensemble des besoins et attentes mentionnées § 5 représente un cas classique. Cependant, il ne faut pas laisser de côté tous les **aspects techniques, organisationnels et humains** *[MERTINA,2003] ; [DESPONDS et al., 1999]* qui se cachent derrière ces volontés. En effet, le choix d'un logiciel peut être influencé par son prix d'achat (licence d'exploitation), avant d'être retenu pour ces qualités. La formation des collaborateurs peut être un frein non négligeable à la volonté de mettre en place un tel système SIG : il est possible que l'organisation actuelle paraisse suffisante pour les collaborateurs qui ne voient alors pas l'intérêt de s'engager dans une voie pour laquelle ils ne perçoivent pas d'avantages directs. Il est donc évident qu'une volonté de la direction doit être manifeste, que les besoins soient validés et approuvés, afin de permettre au projet de mise en œuvre du SIG de voir le jour.
Ainsi, un SIG ne s'improvise pas, il ne se met pas en place non plus grâce à un coup de baguette magique. Il est le résultat d'une **démarche collaboratrice entre acteurs**. C'est ainsi que la collaboration entre le concepteur et les utilisateurs doit être la plus efficace possible. Dans la plupart des cas, les utilisateurs ne connaissent pas les outils de type SIG, leur fonctionnement et leur utilisation. D'autre part, le concepteur ne connaît pas la matière dans laquelle les utilisateurs travaillent au quotidien. C'est pour cette raison que la collaboration, si l'on veut qu'elle soit fructueuse, doit être étroite afin que la compréhension des parties soit parfaite. Souvent une approche pratique des concepts SIG fonctionne mieux qu'une approche théorique.

C'est cette approche qui a été mise en place pour ce travail au sein de la RVC :

- Choisir les moyens techniques en commun, ne rien imposer, aller vers les moyens les plus opérationnels,
- Respecter au mieux l'organisation actuelle des membres de l'entreprise,
- Mettre en place un processus participatif.

Ceci conditionne la réussite du projet et surtout son acquisition « grandeur nature », par extension du prototype.

Résumé des points importants :

La méthode de travail s'inspire d'une démarche conventionnelle en SIG.



Dans ce chapitre :

L'organisation générale de la station de recherche
Des explications concernant le processus de réalisation des essais : le système d'expérimentation..
Le mode de gestion des données.
Réflexion et analyse de ce fonctionnement.

4 Description du fonctionnement actuel à la RVC

On décrit ici le fonctionnement actuel de la RVC *[RAC, 1996]*, *[MERTINA,2003]* ; celui ci est adapté à la situation et à la place de la RVC au sein de la RAC. On rappelle que la RVC **réalise des essais sur les modes de conduite de la vigne,** elle ne réalise pas en principe d'essais phytosanitaires.

4.1 Les projets

Les différents projets ou études menés à la RVC sont des essais (consulté un exemple de travaux :*[ZUFFEREY, 2000]).* Ces essais sont de deux types :

- o Les essais permanents,
- o Les essais temporaires,

Les essais permanents sont des projets à long terme, suivis dans le temps au fil des années. Ils permettent le suivi de la vigne au niveau de ses paramètres les plus essentiels afin de connaître et d'établir par là une base de données de sa « vie ». Leur utilisation à des fins statistiques est majeure, surtout en statistique temporelle. L'exemple le plus simple et le plus courant est celui du suivi de la production en jus de raisin au fil des années.
Les essais temporaires sont des projets de court terme, durant le temps de l'expérimentation et du recueil des données ciblées. En effet, ces projets sont ciblés dans le sens où leur objectif est défini pour un but précis. Ce genre d'étude peut durer de la simple journée d'expérimentation, à plusieurs années de mesures. Ils peuvent déboucher sur toutes sortes de résultats et d'analyses. Une tel essai, serait, dans un cas très court, l'étude de l'effet d'un orage de grêle, et dans un cas plus long, l'étude de l'effet de l'orientation des lignes de vignes sur la maturité du raisin.

Ces essais, quel que soit leur type, sont réalisés sur des parcelles le permettant. On appelle ses parcelles des **parcelles d'essais**. (Cf. fig. 3 : plan du domaine) En effet, selon l'essai mené, le choix de la

parcelle d'essais est différent. Ce choix est réalisé en fonction de la disponibilité des parcelles pour les essais, de la possibilité de réaliser cet essai sur telle parcelle, et de l'expérience des collaborateurs qui connaissent mieux que personne le domaine et ses particularités.
Les projets peuvent être suspendus temporairement pour des raisons diverses, puis repris en temps ultérieur. On dit alors qu'ils sont **en attente**.

De plus, les projets peuvent être réalisés en collaboration avec d'autres institutions, nationales ou internationales : on parle alors de **collaboration externe** dans ce cas, a contrario d'**interne** pour des projets traités au sein de la RVC. Pour exemple nous citerons la collaboration avec un autre département de la RAC elle-même, ou bien encore un partenariat français avec l' Institut National de Recherche Agronomique (INRA)[8].

4.2 Le travail

On peut noter la particularité des chercheurs à travailler en **autonomie**. Ceci s'explique par le fait que ce ne sont pas toujours les mêmes chercheurs qui travaillent sur les mêmes projets, au même moment et avec les mêmes collaborateurs. Des chercheurs peuvent aussi venir momentanément travailler à la station viticole pour un projet temporaire. Ainsi, l'autonomie est de rigueur afin de mener à bien les projets. Ceci permet aussi de toujours avoir une personne **responsable** par projet mais aussi par domaine de recherche.

4.3 Les données

Pour ce qui est des projets permanents, l'organisation est rodée et à fait ses preuves : elle est adaptée à ce mode de fonctionnement. En effet, les données communes sont regroupées, une fois relevées, dans des bases de données. Ensuite, les collaborateurs peuvent s'y référer, réaliser leurs traitements et ajouter leurs données personnelles afin d'obtenir le résultat cherché. Ceci démontre le travail en **autonomie relative**.

Dans tous les cas d'étude, une **multitude de données** sont relevées. Il est important de signaler que **l'usage des données** peut être personnel mais aussi commun. Il y a donc **utilisation multiple** des données.
Les données relevées à la RVC sont issues de son domaine expérimental. (Cf. fig.3 : plan du domaine).
Les autres données peuvent venir d'un lieu d'expérimentation externe, le plus souvent d'autres stations fédérales.
L'ensemble des données relevées n'est **pas référencé** spatialement au terrain. Autrement dit, les données viennent du territoire, qui est en occurrence l'ensemble des parcelles d'essais, mais sans possibilité de les visualiser. Ces données sont alors liées au territoire par des **identifiants** permettant de retrouver leur lieu de provenance. Un « code[9] » est donc en place afin de retrouver le lieu : « **tel code ⇔ tel lieu** ».
Cependant, on note immédiatement la difficulté pour un chercheur « nouveau » de se retrouver rapidement dans tous ces lieux, l'idée de rattacher visuellement ces données au terrain est née : il serait alors plus facile de se repérer, (n'oublions pas que le cerveau mémorise bien plus facilement des données visuelles que des données dites codées !).

---

[8] INRA, Siège : 147, rue de l'Université, 75338 Paris Cedex 07 France.

[9] Un code est un système de symboles, abrégé par rapport au langage ordinaire, représentant une information, ici une localisation sur le terrain.

La figure 4 ci-dessous résume le fonctionnement simplifié de la RVC. Une partie seulement des **cas de figure** est abordée, car des **combinaisons de cas** peuvent avoir lieu, et des cas spéciaux existent, comme toujours dans ce genre de travaux d'expérimentation.

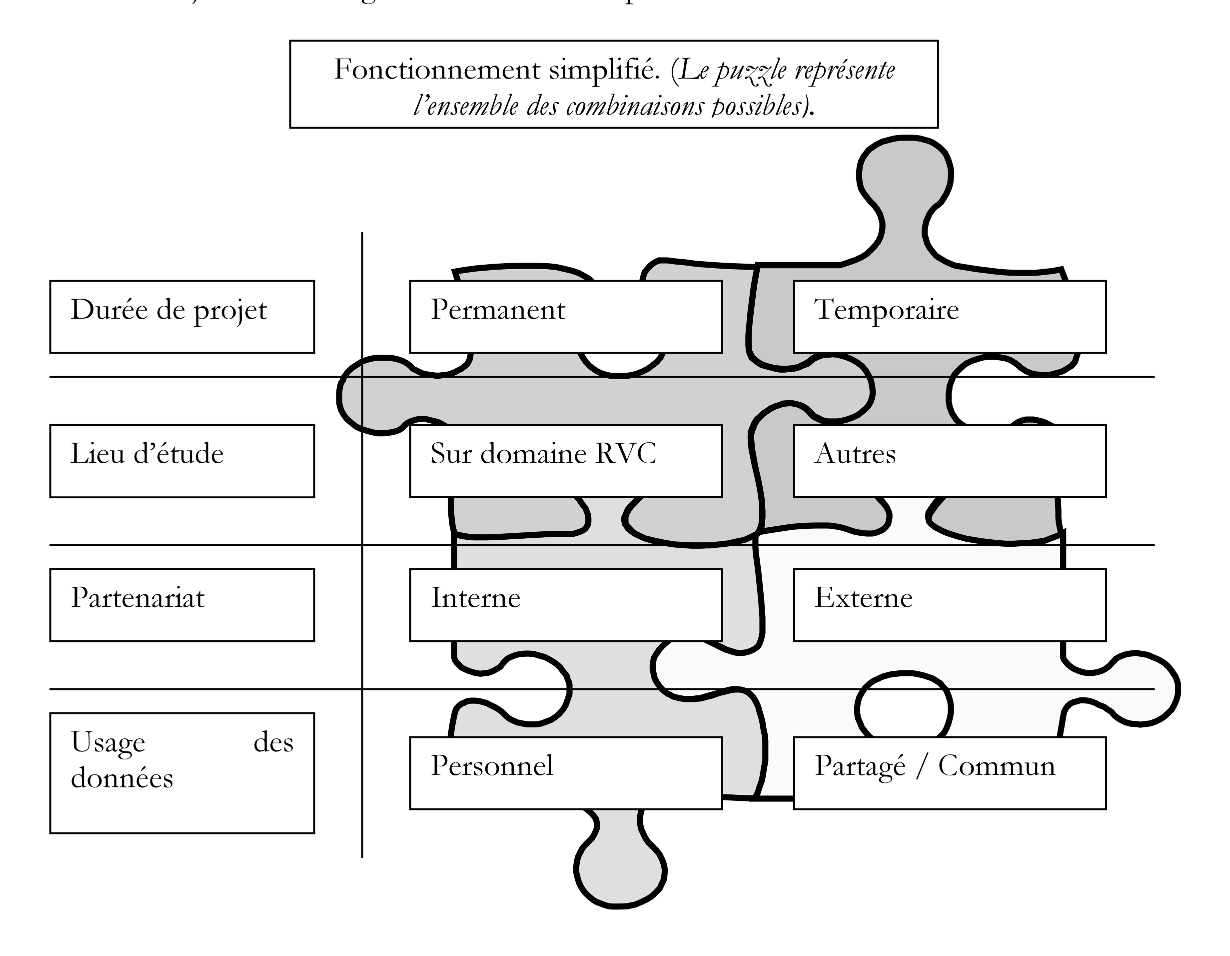


Figure 4 : résumé, sans détail, du fonctionnement de la RVC.

4.4 Le plan

La station possède **un unique plan papier** de ses parcelles d'essais ! L'unique possibilité de se repérer est de faire une photocopie A3 d'un plan réalisé à partir d'un extrait cadastral oú les parcelles internes ont été tracées ¨ à la règle ¨ pour les délimiter. (Cf. fig. 3: plan du domaine expérimental de Caudoz). Il en est de même avec des plans intérieurs des parcelles montrant les lignes de vigne, et la codification adoptée pour référencer les données : ces plans sont des illustrations de la parcelle, réalisés sous Excel ou à la main, ils servent uniquement à se repérer et à prendre quelques notes personnelles. (Cf. fig. 5 exemple de plan intérieur de parcelle d'essai).

4.5 Les cartes

Quelques **cartes du domaine** sont à disposition des chercheurs sous forme « papier » le plus souvent. Ce sont par exemple : la **carte pédologique** des lieux, la **carte de la réserve utile en eau du sol**. Ces cartes sont réalisées par des bureaux externes mandatés. Elles sont imprimées et stockées afin de pouvoir s'en servir pour d'éventuels travaux, notamment des travaux de corrélation entre phénomènes, (exemple : la teneur en eau du sol a-t-elle une influence sur la maturité du raisin ?...).

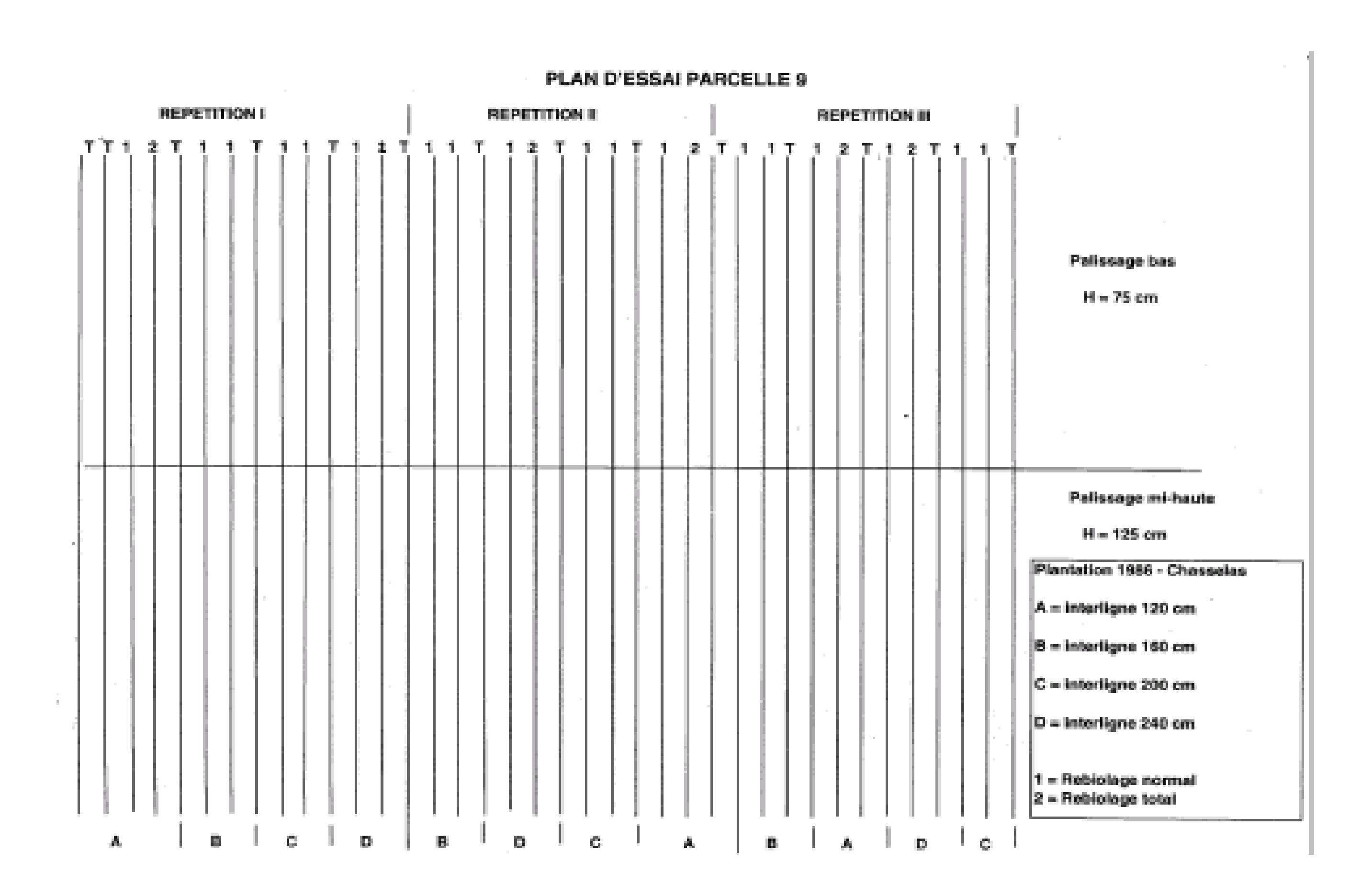


Figure 5 *: exemple de plan intérieur d'une parcelle d'essai.*

4.6 Le système d'expérimentation

4.6.1 Les essais

Le système d'expérimentation se base, pour le vignoble, et en général pour l'agriculture, sur le **système de conduite**.
Le système de conduite est un **mode de culture** choisi pour le vignoble, c'est-à-dire la façon de **mener la vigne**, et ceci se réalise sur le long terme. En effet, il existe plusieurs façons de conduire le vignoble : palissage haut, bas ; interligne large, fin, rebiolage normal, total, , orientation des rangs, hauteur de tronc, système de taille, … etc.
Pour mener ces expérimentations, des **essais** sont réalisés. Ces essais suivent les objectifs des études, ils sont regroupés en 4 thèmes : *amélioration et multiplication du matériel végétal, techniques culturales (entretien des sols et modes de conduite), techniques vinicoles, chimie des vins.* **Un essai représente donc la réalisation concrète d'une expérimentation sur une parcelle.** Il y a toujours un essai par parcelle, qui effectue les tests nécessaires pour l'étude ou le projet en question. **L'essai est la source de données**.

4.6.2 Organisation des essais sur les parcelles

Nous avons vu que des essais sont réalisés sur des parcelles dites **parcelles d'essais**. Ces parcelles d'essais sont alors découpées en une **grille** qui partitionne le territoire selon plusieurs **niveaux,**, elle s'adapte aux essais oú viennent se rattacher les **données**.

Les niveaux constituent à chaque fois une **partition inclusive** du territoire. Ceci signifie que la parcelle est découpée en répétitions, qui contiennent plusieurs variantes composées de lignes elles-mêmes composées de pieds. La table 5 ci-dessous explique l'organisation :

NIVEAU	DESIGNATION	UTILITE
Parcelle	Partition du domaine de la RVC	Permettre la **réalisation d'un essai** répondant aux projets.
Répétition	Subdivision d'une parcelle en plusieurs parties, généralement 3 ou 4.	Permettre une répétition des variantes de l'essai afin d'obtenir un **échantillon statistique** représentatif des données, avec absence de biais ou d'hétérogénéité du terrain d'essai.
Variante ou procédé	Représente la problématique étudiée. Elle répétée sur plusieurs répétitions.	Permettre la réalisation de **différentes variantes pour un essai.**
Ligne	Division de la variante au nombre de lignes la composant	Permettre la réalisation d'une variante **homogène, identique** sur une même ligne,
Pied	Décomposition de la ligne au nombre de pieds la composant (cas exceptionnel)	Connaître **l'origine exacte des données** issues d'un essai, à l'entité d'étude la plus fine.

Table 6 : les niveaux du système d'expérimentation, leur désignation et leur utilité.

Ainsi, il y aura, normalement, autant de vins différents générés que de variantes d'essais, pour une parcelle. Ces vins sont issus des mêmes modes de conduite, ils sont donc de même nature. On omet donc souvent de préciser quelles sont les répétitions (justes utiles pour le brassage statistique des données), et les lignes (le détail de mesure des données n'est pas nécessaire, mais toutefois on retrouve de quelle ligne proviennent les données car il y a unicité de la variante au sein de la répétition) pour se contenter au **repérage** suivant :

domaine, parcelle, titre de l'essai et variante considérée.

Exemple : Domaine de Pully, parcelle 20, essai hauteur de tronc, variante Ba

Pour **résumé**, on peut dire que :

- chaque domaine réalise des essais,
- on réalise un essai par parcelle,
- chaque essai est composé de plusieurs variantes,
- on répète les variantes plusieurs fois sur les répétitions,
- chaque variante se réalise sur des lignes de vignes,
- une ligne est composée de plusieurs ceps homogènes.

Ces niveaux sont **codés** de sorte à pourvoir « suivre » les données :

Pour telle date :

Tel code ⇔ tel essais⇔ tel lieu d'expérimentation ⇔ telles données relevées :

Le codage adopté et validé au fil des années est le suivant : table 6 . L'identifiant pour la RVC est le chiffre 1. (respectivement 2, 3 ,4 pour les autres stations rattachées à la RAC).

NIVEAU	CODE	EXEMPLE
Parcelle	de 1 à 66, non continu.	9
Répétition	Chiffres romains ou arabes	II ou 2
Variante	Lettre majuscule, de A à D en général	A : interligne[10] 120 cm B : interligne 160 cm C : Interligne 200 cm D : interligne 240 cm. H : palissage[11] mi-haut h : palissage bas 1 : rebiolage[12] normal 2 : rebiolage total
S'il y a Combinaison de variantes	Codage spécifique selon le système de conduite représenté sur la ligne. Généralement une lettre associée à un chiffre.	AH1
ligne	Numéro croissant, Sauf pour les lignes T signifiant tampon[13]	13, 14, 15, 16, T,T , 19, 20 …
Pied	Nombre entier, numérotation par ordre logique selon les cas.	12, 13, 14, 15…

Table 7 : le codage du système d'expérimentation.

Il faut aussi mentionner que l'usage de la **date ou de l'année d'essai** est indispensable, afin de ne pas mélanger les données au fil des années !

Un **exemple de localisation** pour le recueil de données pourrait être le suivant :

92Dh2, pour l'année 2000 Parcelle 9, répétition 2, variante Dh2

[10] Interligne écartement des rangs, à ne pas confondre avec l'espace intercep, qui représente l'espace entre chaque pied de vigne sur une même ligne.
[11] Palissage : opération qui consiste à attacher les branches de la vigne à un support pour les faire pousser en espalier, c'est à dire superposées dans un plan verticale.
[12] Rebiolage : opération qui consiste à supprimer les entre-cœurs,dits rebiots. Un rebiolage total correspond à la suppression de tous les entre-cœurs, alors qu'un rebiolage normal consiste à la suppression des rebiots uniquement dans la zone des grappes. .
[13] Ligne tampon : ligne de pieds de vignes permettant de séparer les essais entre eux. En général chaque variante est séparée par une ou plusieurs lignes tampons (T).

Equivalant à : Parcelle 9, répétition 2, variante D, palissage h, rebiolage 2

Ou bien encore : Parcelle 9, répétition 2, interligne 240, palissage bas, rebiolage total.

La localisation du pied n'est pas nécessaire dans la plupart des cas, sauf **applications spéciales**, du fait que la ligne est composée d'un ensemble homogène de mêmes pieds.

La figure 6 montre l'exemple de ci-dessus. Le trait épais localise la chose.

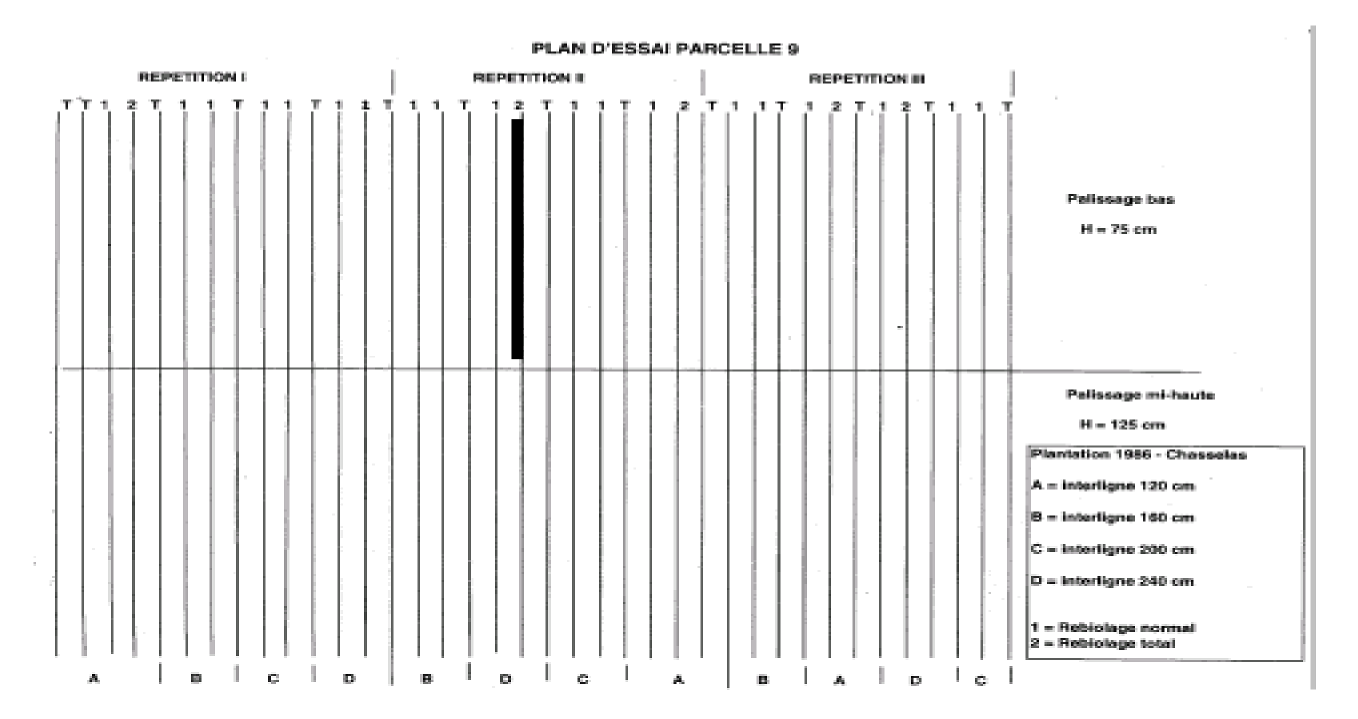


Figure 6 : exemple de localisation de 92Dh2, sur une parcelle d'essai.

4.6.3 Utilité de ce mode d'organisation des essais

Ce mode d'organisation des essais permet :

- D'avoir à disposition les **données d'essais**,
- De pouvoir **s'adapter aux besoins de l'expérience en cours d'expérimentation**,
- De pouvoir **changer d'essais** lorsqu'il en est question
- D'obtenir **un échantillon représentatif des données** grâce à un « brassage » de celles –ci sur le territoire (grâce à la répétition) permettant ainsi de procéder à des **analyses statistiques**. (En d'autres termes, il y a répétition des essais afin de ne pas introduire de biais ou de fausser des résultats.)
- De créer un **historique** complet lorsque les données sont assorties de leur date. (Cependant, la location spatiale de « l'ancienne » parcelle d'essai n'est pas possible).

Ainsi, ce mode d'organisation est totalement adapté aux expérimentations, sa flexibilité le rend pleinement opérationnel, à condition de toujours conserver **le code et la date associée**.

Ce mode d'organisation n'est **pas unique**. Ce système d'expérimentation peut s'appliquer aussi aux autres domaines de recherche fonctionnant selon le même principe : la parcelle d'essais, la répétition des essais, les variantes étudiées, les lignes d'expériences, les pieds d'études.

4.7 La gestion des données

4.7.1 Cas général

Chaque donnée du domaine est structurée selon son appartenance à un thème. Pour chaque thème, il y a édition d'un **cahier** annuel comportant toutes les données s'y référant : ce sont des **bases de données thématiques** :

- Données de base pour chaque élément du système d'expérimentation,
- Données de vinification,
- Données de dégustation,
- Données de contrôle
- Données de vendanges,
- Données physiologiques,

La plupart de ces cahiers sont édités sur Excel ou Word, ils présentent les données mais aussi des résultats, des bilans… ensuite, ces cahiers ont la possibilité d'être comparés inter annuellement et inter cahier afin de créer des bilans et des résultats principaux. Le mode de fonctionnement dit en **indépendance** plus haut**,** se justifie par le fait que l'accès, la consultation, la mise en œuvre … des bases de données thématiques est libre. Chaque collaborateur a à sa disposition les cahiers et ses **propres données personnelles** qu'il peut utiliser selon ses **volontés**.

Généralement, les données des projets temporaires sont individuelles et personnelles, celles des projets permanents sont des **données communes.**

Les données d'intérêt sont des données dont chaque collaborateur a généralement besoin de connaître au sein d'une base de données thématiques.

Ainsi, le futur SIG ne pourra pas intégrer, pour ce présent travail, l'ensemble des données, mais les données communes et d'intérêt.

Une fois que toutes les données d'étude ont été acquises, alors elles sont rassemblées, traitées et publiées. Il est important de signaler que toutes ces données sont internes au domaine, elles ne sont pas diffusées directement de manière brute mais sous forme de synthèses, par peur de mauvaise interprétation et d'impossibilité d'exploitation.

La figure 7 présente le diagramme conceptuel d'activité simplifiée pour la RVC.

La figure 8 présente le diagramme de flux de données, pour la partie qui nous intéresse dans ce travail.

Le § 7 explique plus précisément les bases de données thématiques, leur contenu et leur structure, dans la partie « inventaire des données ».

DIAGRAMME CONCEPTUEL D'ACTIVITE SIMPLIFIE (DCA)

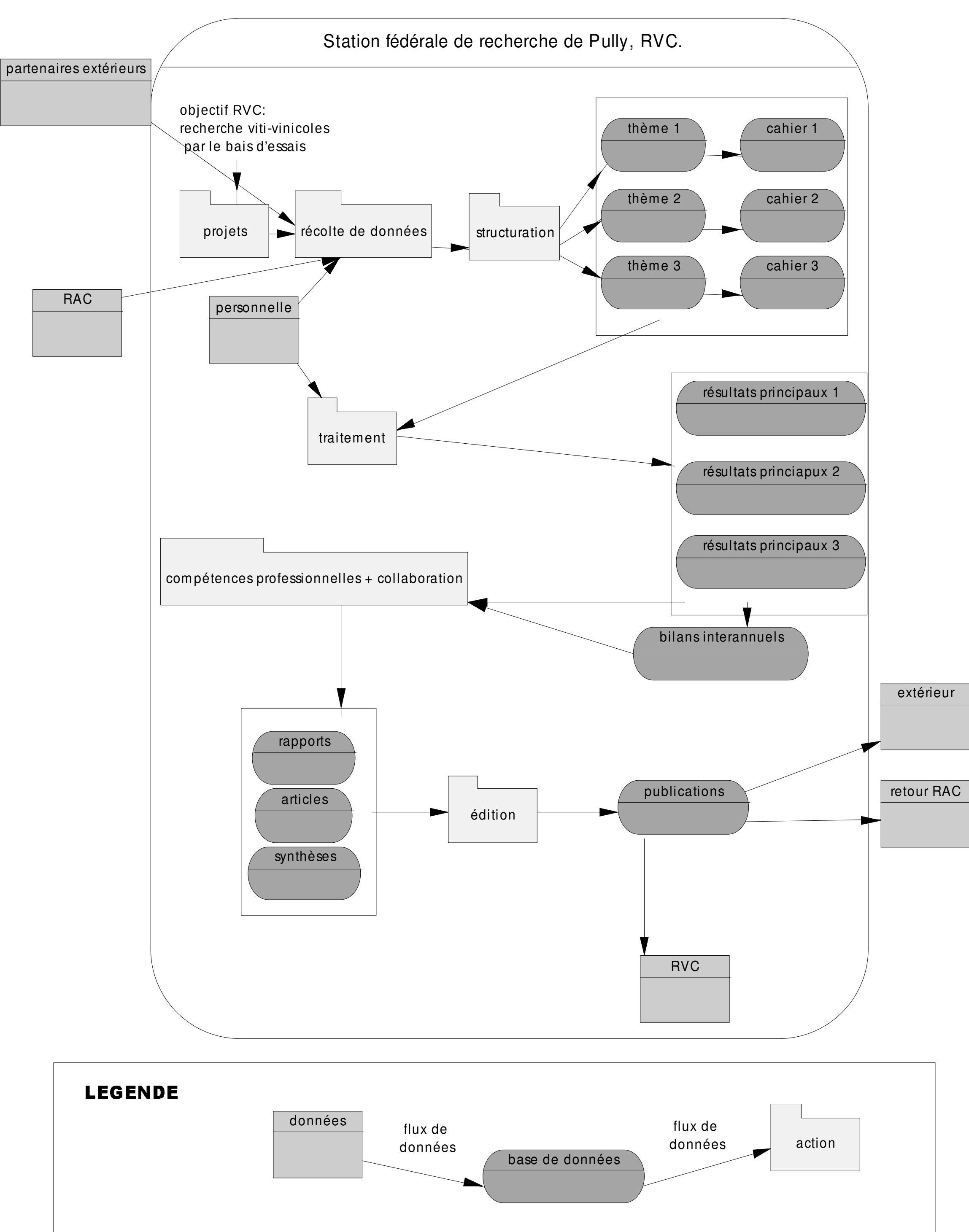
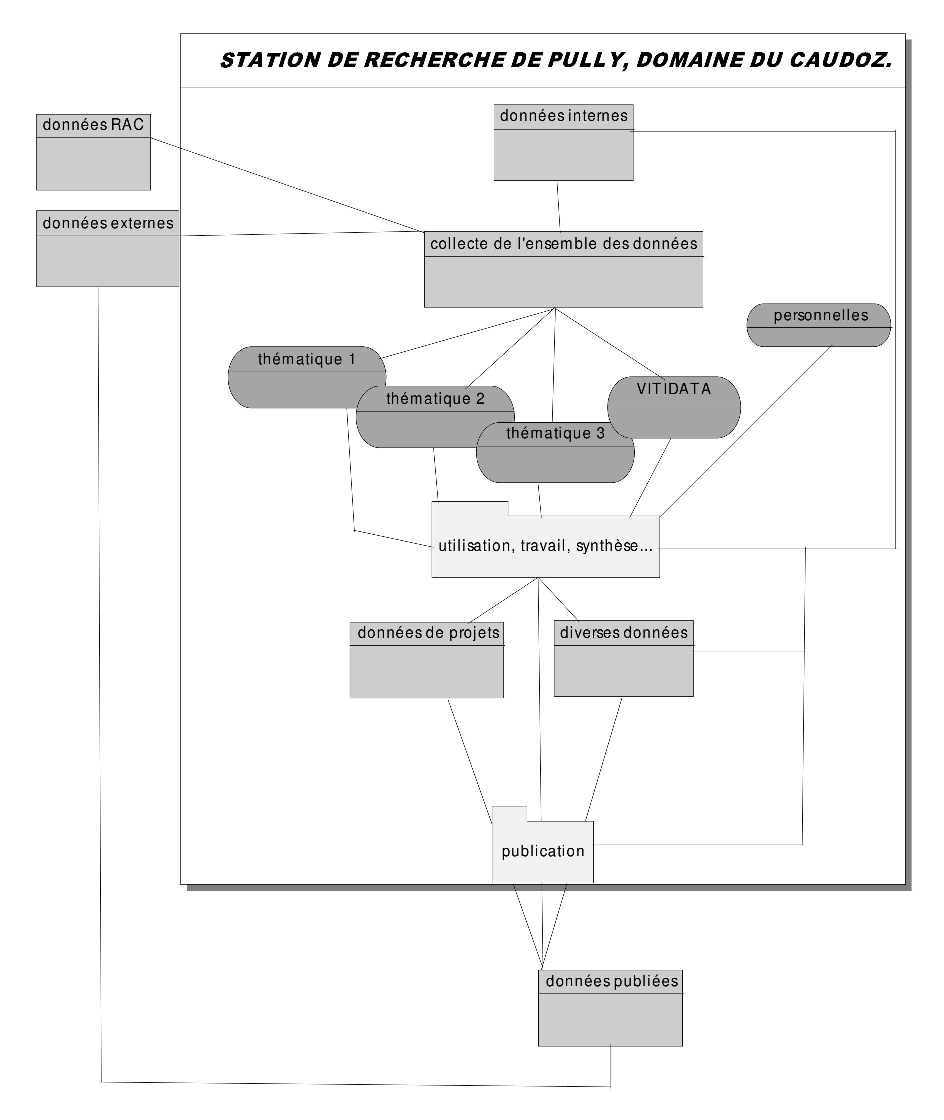


Figure 7 : diagramme conceptuel d'activité de la RVC. (Simplifié pour le présent travail).

Diagramme simplifié de Flux de Données (DFD)



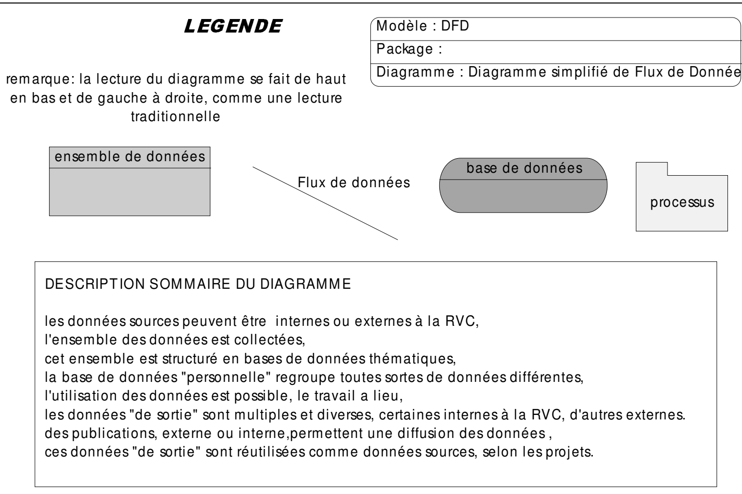


Figure 8 : diagramme de flux de données de la RVC. (Simplifié pour le présent travail).

4.7.2 Un cas particulier : VITIDATA

Les données thématiques « paramètres viticoles » sont centralisées, une fois mesurées, grâce à une base de données principale nommée **VITIDATA[14],** d'oú résulte alors le **cahier de vendange** (Cf. table. 7 le système VITIDATA existant (contenu et code) et § 7.2-3 : le cahier de vendange). Le codage permettant de lier les données aux essais et par conséquent au terrain, est basé sur le système d'expérimentation. Toutefois, VITIDATA ne code pas la répétition mais elle ajoute un identifiant de type entier (pour simple moyen de recherche rapide) ainsi que le numéro de la station de recherche, soit pour nous le numéro 1. Ceci se réalise sous le logiciel Access. Une édition de codes barres est possible par VITIDATA : ils sont apposés sur les caisses de vendange et permettent ainsi d'identifier automatiquement la provenance et d'y faire correspondre les données s'y référant, ceci lors de la vinification au laboratoire.

CONTENU DE VITIDATA

Données générales	**Exemple**
Domaine	Changin
Parcelle	5
Cépage	Aligoté
Interligne	200/115
Plantation	1984
Conduite	Guyot simple
Porte-greffe	3309
Intercep	85
Tampon	Gewürztraminer
Données de vendange	**Exemple**
Répétition	A1
Index	2.5.43
Surface	21.44
Nombre de ceps	16
Clone	263
Date de vendange	28.09
Kilogramme	24.4
Kg/m2	1.138
IR Brix	21.6
IR °Oe	90.1
D20/20	1.0924
PH	3.10
Acide total	8.8
Acide tartrique	8.0
Acide malique	3.1
Ind. form.	11.0

EXEMPLES DE CODES D'ESSAIS
Essai simple : 2.5.43.A1, pour : domaine de Changin, parcelle 5, identifiant 43, variante A, ligne 1.
Essai complexe : 1.27.1.A,NC,NS,non rebiolé, pour : domaine de Caudoz, parcelle 27, identifiant 1, variante A, ligne NC,NS, non rebiolé.

Table 8 : les données de VITIDATA et exemple de codes d'essais

[14] VITIDATA a été mise en place en 1998 par un bureau privé.

4.7.3 Le traitement des données

Le **traitement des données** se fait surtout de manière individuelle sur la base des données communes des différentes bases thématiques et aussi des données individuelles mesurées. De là, sont tirés les véritables **résultats des études**. Les traitements sont le plus souvent des analyses, des moyennes, des tests statistiques…ou résultats de fonctions.

4.7.4 La publication des données

Les résultats, conclusions des études sont publiés (exemples : *[ZUFFEREY,2000], [RAC,1996])* dans des **rapports officiels**, mais aussi sous forme **d'articles de synthèse** dans des revues spécialisées telle que *la revue suisse d''agriculture*. C'est ici l'unique accès aux données disponibles pour le public : sous forme de conclusions, sans accès direct aux données mesurées, qui sont conservées à des fins statistiques et historiques.

Le § 7 traite plus en détail cette partie « données ».

4.7.5 Mise à jour des données d'essais

La mise à jour des données consiste au remplacement d'une donnée par une autre, l'ancienne n'étant plus d'actualité, c'est à dire obsolète.
Pour la RVC, le problème ne se pose pas. **La mise à jour des données d'essais n'est pas nécessaire**. En effet, les données sont stables. Une fois relevées, elles restent fixes. Ceci est dû au fait que ces données sont datées, il y a donc notion d**'historique**.
Mais attention : il y a juste mise à jour lorsque l'essai parcellaire change : il y a alors naissance d'une nouvelle parcelle d'essais à la place de l'ancienne.
Toutefois, les données non relatives aux essais suivent une mise à jour classique : lorsqu'une donnée n'est plus d'actualité, on l'efface et on la remplace par la nouvelle, valable. C'est le cas pour les données générales administratives.

4.7.6 Historique et traçabilité

La gestion de l'historique est assurée par la **date**, souvent remplacée par l'année d'essai. La **traçabilité** (voir à ce sujet le § 10) concerne le suivi d'un élément de son lieu de production à son lieu de consommation. Ici, **de la vigne à la bouteille de vin**. Afin de permettre cette traçabilité, il y a des codes (ou identifiants), différents selon chaque base de données thématiques, et donc aussi selon chaque chercheur pour les données personnelles, qui permettent de « **remonter les informations** ». Ceci se réalise par le titre de l'essai et le nom du domaine que chaque chercheur connaît et retrouve !
Il n'y a pas **d'harmonie entre codification** des différentes bases de données thématiques. Le futur SIG ne devrait apporter plus qu'un seul code : l'identifiant unique de l'élément d'essai. Le SIG devra regrouper ces différents codes afin d'accéder aux autres données.

4.7.7 Normalisation

Au sein d'une base de données, donc aussi d'un SIG, une multitude de données peut être enregistrée. Afin que ces données soient **cohérentes**, elles doivent être **normées**. Ceci signifie qu'elles doivent répondre à des critères précis permettant la cohérence entre elles. Il faut ainsi :

- un **dictionnaire des données**, permettant de connaître les données, leur type, leur représentation…
- des **méta données**, permettant d'obtenir des données sur les données, c'est à dire des sources d'information sur les données,

- des **normes internes de fonctionnement**, comme par exemple des indications sur les formats d'échange des données, leur structuration, leur responsable…

Au sein de la RVC, la normalisation est partielle. On distingue les données normalisées : ce sont celles qui entrent dans les bases de données thématiques. Il y aussi les autres données qui sont personnelles et qui sont donc ainsi au goût de son prétendant.
Ainsi, on peut dire que les dictionnaires de données sont absents car les chercheurs disposent de ces dictionnaires en tant que **connaissance professionnelle**, si il y a risque de confusion (unité…) ceci est mentionné. Les **méta données** apparaissent sous les termes « date, personne, matériel et méthode » au sein de chaque base de données, conservées par son responsable.
Pour le futur SIG, un **dictionnaire descriptif de la base de données** sera réalisé comme il se doit.

4.7.8 Sauvegarde et sécurité des données

La sécurité des données est assurée par des sauvegardes. Celles-ci sont effectuées quotidiennement par les services en charge, au niveau de la confédération (le gestionnaire). Cela revient à une garantie de sauvegarde des données enregistrées sur les comptes utilisateurs des collaborateurs. Il ne va pas sans dire que chaque personne est responsable de ses données en dehors des responsabilités du gestionnaire.

4.8 Analyse de ce fonctionnement

4.8.1 Les difficultés à considérer pour le futur SIG dont le problème spatio-temporel

Ce mode de fonctionnement comporte quelques difficultés qu'il faudra maîtriser :

- **Limites de parcelles non fixes selon les essais réalisés** : ceci représente un cas SPATIO TEMPOREL. En effet, les parcelles peuvent se voir augmenter ou au contraire diminuer de quelques lignes de vignes selon les études, notamment au niveau de leur nombre de lignes faisant office de tampon (lignes T), elles peuvent en outre subir toute sorte de transformation selon le bon vouloir des chercheurs pour les besoins de leurs expérimentations. Cette particularité pourra être surmontée par la séparation de l'objet de ses données. (CF.§ 9 : réalisation du prototype).
- **Réaliser la différence entre la parcelle physique (l'objet) et la parcelle d'essai (la table contenant les données)**: l'une étant la parcelle du plan et l'autre étant la concrétisation pour une durée définie de la réalisation d'un essai.
- **Une gestion de l'historique importante** : comme la plupart des données ont une relation temporelle, l'accès aux données antérieures devra être possible.
- **La notion de traçabilité** : elle devra être précisée et sa valeur dans un tel contexte réévaluée afin d'uniformiser sa gestion
- **Un modèle de données interne au SIG se devant d'aller vers les données prioritaires** : le SIG devrait être réalisé d'abord pour les données appartenant aux projets permanents puis au cas pas cas, pour les projets temporaires. Les données communes et d'intérêt devront être traitées en priorité.
- **Les redondances** : la structuration des données au sein d'un modèle commun devra les éviter,
- **Les contraintes techniques, organisationnelles et humaines** : elles pourront influencer la mise en œuvre du SIG de manière considérable si des réelles volontés de progrès sont absentes de la part des collaborateurs.

- **Les normalisations** : devront être effectuées pour le SIG grâce à un dictionnaire, sans changement manifeste pour les collaborateurs,
- **La volonté de ne pas chambouler le fonctionnement actuel** : ceci pour permettre au SIG d'être un moyen additionnel en tant qu'outil nouveau (Cf. § 6 : perspectives du futur SIG).
- **Veiller à l'intégration du SIG au sein de la RVC** : ceci afin de permettre son utilisation judicieuse.

4.8.2 Les faiblesses de ce fonctionnement

Le tableau 9 ci-dessous, illustre le fait que ce mode de fonctionnement comporte des risques pour le monde des SIG, que contourne d'une certaine manière la RVC.

FAIBLESSE	POURQUOI ?	SOLUTION RVC
Indépendance	. Il y a risque de duplication, de redondance et un usage de données supposées communes mais en réalité différentes pourrait se présenter.	L'indépendance est adaptée aux projets de recherche, ceci n'excluT pas une collaboration entre chercheurs. Les données sont issues de bases communes avant toute utilisation. Les projets sont bien répartis, chaque personne sait ce qu'elle a à faire.
Normalisation	Si il n'y a pas de dictionnaire décrivant les données, et que les méta-données sont absentes, alors il y a risque de confusion et d'emploi de données non plausibles. La simple consultation de données rassemblées par un autre collaborateur devient vite difficile, sauf si ce dernier a pris soin de documenter son travail	Les données habituelles sont acquises suite à la formation. Les méta données sont relevées par le chercheur en charge du projet sous la forme : date, personne, lieu, matériel, méthode.
Centralisation	Le fonctionnement actuel ne permet pas un regroupement commun des données, ainsi, il est difficile de parvenir à un renseignement se trouvant dans une « autre base. »	La centralisation des données n'a pas lieu. Seules les bases de données thématiques existent. Chaque chercheur se sert de ce qu'il a besoin parmi une panoplie de données. Il dispose en outre de ses propres données.
Restriction de potentialité	Que cela soit pour la visualisation, pour la production de cartes ou l'édition de documents, les possibilités actuelles sont faibles voire inexistantes.	Les possibilités de traitement sont effectuées grâce à des logiciels fédéraux identiques pour toutes les stations de recherche (WIDAS). Pour l'édition cartographique, les possibilités sont absentes.
Structuration des données	Malgré la multitude et la diversité des données, celles ci ne sont pas toujours structurées convenablement. Ceci est renforcé par l'individualisme régnant.	La structuration des données est en ordre pour les bases communes, mais sinon chaque chercheur est responsable de ses données, structurées ou pas.
Dispersion des données	Il y a absence de système fédérateur. La codification n'est pas la même selon la base thématique, ceci crée alors une recherche des informations difficiles si ce ne sont pas les siennes.	Sauf pour les bases thématiques communes, le restant des données est dispersé au sein de chaque chercheur.
Communication	Améliorer le pôle communication en utilisant des supports modernes de diffusion, créer encore une fois l'accès au terrain par des cartes thématiques illustrant les projets, les conclusions…	Diffusion sous d'autres formes, réalisation avec d'autres logiciels sous d'autres formes (tableur, graphiques…), ,mais réel besoin d'accès au domaine !

Table 9 : les points dangereux du fonctionnement de la RVC, leur solution au sein de la station.

On remarque que le fonctionnement actuel n'est tout simplement pas adapté à une gestion commune des données, il y a à chaque fois une parade adaptée qui contourne les points dangereux pour les SIG. D'ailleurs cette gestion commune des données est –elle possible par un SIG pour une entreprise si complexe ?. En effet, ce mode de fonctionnement est au contraire parfait pour des personnes travaillant en indépendance à partir de données regroupées par thème, ce qui est le cas. .

Cependant, notre rôle n'est pas de **chambouler une organisation** et un mode de fonctionnement acquis, mais nous nous devons de remarquer les dysfonctionnements, les problèmes… de manière à en **tenir compte lors de la mise en œuvre du SIG**, qui se devra d'être adapté à la situation, et répondant aux besoins et attentes.
Ainsi, les besoins sont réels, l'apport est indéniable. Si comme souvent « l'offre crée le besoin »…alors le SIG aura sa place à la RVC dont il s'agit encore de définir les rôles. (Cf. §6 perspectives de développement)

Résumé des points importants :

Le fonctionnement de la RVC est complexe, il est le résultat d'années d'expérience de la part des chercheurs. Le travail en indépendance est le mode d'organisation qui a fait ses preuves. Ainsi, le SIG doit tenir compte de ce fonctionnement particulier, sans le chambouler, mais plutôt en s'adaptant, en apportant des choses nouvelles et en créant de nouvelles perspectives.
Le SIG devra se concentrer sur les données dites communes et d'intérêt afin d'être viable.
Le système d'expérimentation mis en place n'est pas unique. Ainsi le SIG pourra fonctionner pour toutes entreprises fonctionnant selon ce principe, en changeant uniquement la structure des tables pour les adapter aux nouveaux cas.



Dans ce chapitre :

Toutes les informations concernant les besoins et attentes formulés par les chercheurs en référence au futur SIG.

5 Evaluation des besoins et des attentes[15]

Ces besoins et attentes sont formulés par les chercheurs de la RVC..
Nous parlons ici de besoins et d'attentes car il y a une différence entre un besoin et une attente. Un besoin est une **nécessité**, alors que une attente est plus considérée comme un **souhait.**
Il est important de signaler que les besoins et attentes donneront naissance à une panoplie d'offres. Cependant, il ne faut jamais oublier que **l'offre crée le besoin** : c'est à dire que lorsque l'on a une offre on s'interroge, on réfléchit et on donne ainsi naissance à des besoins supplémentaires nouveaux, auxquels nous n'avions pas pensé auparavant.
Ces besoins et attentes peuvent avoir lieu au niveau des données (consultation, visualisation) mais aussi au niveau de leur utilisation (gestion, mise en valeur et traitement). [*MERTINA, 2003*].

La fig. 11 récapitule ces besoins.

5.1 Besoins / attentes en consultations

Les données étant nombreuses et diversifiées, il ressort que la **simple consultation de données** en tant qu'**information** est indispensable. Ces informations pourront être utilisées directement ou bien pour joindre en complément à un projet. Certaines seront référencées au domaine, d'autres non. Les données consultées devraient être en premier lieu les **données d'intérêt ou / et communes**. On veut dire par-là que la consultation précise d'une donnée ciblée ne devrait pas être nécessaire, mais au contraire, une moyenne de données généralisées serait quant à elle indispensable. La référence au terrain lors de la consultation permettrait de se localiser et d'obtenir ainsi des données issues d'autres parcelles d'essais permettant **l'accès aux données de collègues**.
Un exemple typique de consultation attendu est celui des données de base : date de plantation, surface de la parcelle, nombre de ceps, type de cépage, etc.…

5.2 Besoins / attentes en visualisations

N'ayant pas de plans corrects à disposition, le SIG devrait permettre **d'éditer des plans** du domaine mais aussi des parcelles, voire des zooms plus précis. Ceci pour l'unique besoin de visualisation et de

[15] Réalisé suite à des entretiens.

prise de notes personnelles, de discussions interactives, de repérage pour les invités et les stagiaires, et de guide de travail pour les ouvriers. La composante est donc indispensable. On retiendra aussi l'ensemble des couches sous forme de cartes telles que la carte pédologique, la carte des épaisseurs de sol (à venir), la carte des teneurs en eau du sol… qui représente des sources **d'informations spatiales**.
Un exemple typique de visualisation est de pouvoir consulter n'importe quel plan de parcelle rapidement, de pouvoir l'éditer, et localiser un endroit. L'introduction de « labels » devrait permettre de localiser les objets rapidement.

5.3 Besoins / attentes en gestions / fonctionnements

Avoir accès aux **fonctions classiques des bases de données** (requêtes, gestion du temps, conservation, archivage des données,…) est demandé. L'identification grâce à la date permettra un historique comme c'est déjà le cas actuellement. Ce que l'on entend par « gestion » signifie plutôt « pouvoir réaliser les manœuvres nécessaires au fonctionnement de l'entreprise » ; donc **gérer l'entreprise**.
Un exemple typique en matière de gestion de données est de pouvoir réaliser une **requête** (interroger le système) afin de retrouver une ancienne donnée afin d'élaborer un suivi dans le temps.

5.4 Besoins / attentes de mises en valeur

Par mise en valeur, on entend ici la création de **cartes thématiques** selon les demandes, l'édition de **documents modernes de diffusion** aussi bien de cartes que de listing, la représentation spatiale de données d'intérêt choisies. Il s'agit donc bien d'un besoin venant se greffer sur les données et non pas créant des données. Ceci apparaîtrait utile pour deux choses principales : l'une est la création de documents de synthèses pour les rapports, l'autres est la création de documents de travail à incidence spatiale.
Un exemple typique de cartes thématiques pouvant être créées est celui des cartes représentant les parcelles en fonction de leur encépagement, de leur espace interligne, de l'enherbement, ou de leur entretien effectué dont les traitements subits.

5.5 Besoins / attentes en traitements

Par opposition au terme « mise en valeur », un traitement de données crée d'autres données par application de fonctions. Un traitement peut être d'appliquer une moyenne sur un échantillon de données semblables. Pour ce qui est de la RVC, les traitements proprement dits sont réalisés grâce à des logiciels précis et définis pour l'ensemble des stations de recherche fédérale : Exemple du programme statistique WIDAS. Ainsi les chercheurs souhaiteraient effectuer des traitements sous forme **d'analyse spatiale,** jusque là impossible sous cette forme, plutôt que d'avoir accès à des fonctions existant déjà avec des logiciels communs. La recherche de corrélation entre la réserve hydrique modale et le poids des bois de taille en est un exemple : voir §10.1.8.
Les exemples restent encore à définir plus précisément en fonction de l'offre du futur SIG et des projets pour lesquels une analyse spatiale sera la bien venue. Toutefois, on peut citer l'exemple d'analyses de corrélation d'un paramètre en fonction des cartes thématiques crées et des cartes spatiales de base comme la pédologie des lieux.

5.6 Besoins / attentes non formulés

Ne sont pas formulés les besoins en terme **d'applications spéciales**. On entend par-là, la modélisation, l'adjonction de modules supplémentaires pour la réalisation d'applications précises, ou encore la composante 3 D. Nous restons ainsi dans le cas d'un **logiciel SIG standard**.

Les chercheurs ne souhaiteraient pas qu'il y ait redondance totale entre les différentes bases de données. Ainsi **le futur SIG se devrait d'être principalement une porte d'entrée à la dimension spatiale, absente actuellement.**

BESOINS	OFFRES	EXEMPLES
CONSULTATION	Accès à l'information simple	Consultation des données de base (cépage, plantation…)
VISUALISATION	Accès à la dimension spatiale	Visualisation des plans du domaine et des parcelles d'essai. Visualisation des couches d'informations.
GESTION	Augmentation du fonctionnement de l'entreprise	Réalisation d'interrogations sur les données de vendange.
MISE EN VALEUR	Création de documents géographiques et de listing	Publication de cartes thématiques et de listes associées.
TRAITEMENT	Fonctions d'analyses spatiales principalement.	Analyse de corrélation entre différentes couches.

Figure 9 *: les besoins / attentes au sein de la RVC et les offres produites.*

Résumé des points importants :

Les besoins et attentes des chercheurs en matière de gestion de données nécessaires au fonctionnement de l'entreprise sont traditionnels pour les SIG. Toutefois, les chercheurs



ne souhaitent pas avoir une redondance totale des données avec les bases thématiques.

La fonction spatiale doit être prioritaire pour des données d'intérêt et communes.

Dans ce chapitre :

Les choix et volontés quant au futur fonctionnement du SIG.

6 Perspectives de fonctionnement du SIG

6.1 Comment voir le futur SIG : choix et volonté

Comme déjà discuter au fil de ce rapport, la proposition est la suivante **:**

Un SIG fédérateur des données essentielles. (intérêt + commune + pour essais redondants et permanents)

Il est important de signaler dès à présent que le choix de ne pas chambouler ce fonctionnement a été réalisé. Le futur SIG devra **s'intégrer dans une structure existante** et répondre aux besoins et attentes des chercheurs de sorte à leur apporter un plus plutôt que des contraintes organisationnelles qui nuiraient à leurs travaux.
Afin de ne pas créer de mauvaises interprétations des données, le SIG sera en accès unique pour les chercheurs et ayants droits, une connexion à l'Internet pour la consultation simple n'est pas souhaitée.

Quelques volontés :

- Possibilité d'extension à l'ensemble des stations de recherche et de leurs sections de recherches utilisant sur les mêmes méthodes d'essais (parcelle, répétition, variante…)
- Possibilité d'être modulable en fonction des projets de recherches,
- Applicable aux autres stations en vue d'une semi-centralisation (pour les données d'intérêts et/ou communes.
- Laisser la structure actuelle afin de ne pas modifier les habitudes de travail dans une organisation déjà passablement complexe,

- Limiter les mélanges des données, rester fidèle au modèle actuel de classement et structuration des données (d'oú la mise en place des blocs de données § suivant)
- Permettre d'être opérationnel rapidement pour utiliser les fonctionnalités nouvelles du SIG, s'exercer et créer leurs nouvelles offres et les nouveaux besoins.
- Répondre précisément aux besoins et attentes des chercheurs : ne pas s'éloigner des objectifs fixer.

6.2 Un système SIG additionnel, complémentaire et modulable

Le SIG dit **additionnel, complémentaire et modulable** est alors la solution adaptée à la situation particulière de la station de recherche.
Par **additionnel** il faut comprendre que le SIG ne doit pas écraser l'existant mais bien le complémenter. Ainsi la base de données VITIDATA doit rester active, les autres bases de données thématiques peuvent, elles aussi, rester actives ou bien disparaître pour être implantées dans le SIG.
Par **modulable**, il faut comprendre que le SIG devrait pouvoir s'adapter facilement aux différents besoins des chercheurs en matière de données et de traitements possibles, ainsi que pour le genre d'essais menés sur le système d'expérimentation. (Notamment les possibilités **spatio-temporelles**).

Le SIG serait alors décomposé en deux parties distinctes :

Une partie permanente servant à la gestion des données courantes, telles que les données issues de la vinification, les données de dégustation, les contrôles et les données dites externes. (Cf. § 7). Ces données seraient des données fixes, relevées selon une norme, permanentes, identiques d'année en année…. Elles forment alors les données de base. (CF.§7 : les blocs de données 2 3 7 4 5 6). Rappelons que ces données sont communes aux projets dans leur ensemble ou du moins que l'ensemble des chercheurs peut en avoir besoin.
La deuxième partie concerne toutes les **données non permanentes** (CF §7 : blocs 7 et 8 illustrés pour nous,) ou bien spéciales à une étude. Elles ne concernent que le chercheur en poste sur l'étude. Ainsi, ces données sont individuelles à la différence des premières plutôt de base. Ces données pourraient être introduites dans le SIG mais ceci grâce à l'intervention d'une personne maîtrisant la structure du logiciel utilisé pour le SIG. Ceci ne devrait pas poser de problème majeur car la base de données VITIDATA ne fonctionne pas en autonomie mais est gérée par un consultant, ces services sont donc de rigueur lors de la période de vinification. Le mandat pourrait être étendu à ce **gestionnaire** pour le SIG et ses applications spéciales mais aussi en général pour la maintenance et le suivi des éventuels problèmes du SIG.

Ces deux parties forment alors un SIG qui répond aux besoins et aux attentes des chercheurs. De plus le fonctionnement actuel n'est pas modifié de manière considérable. La solution est **modulable**, et complémentaire au système actuel. Le SIG apporterait alors une touche de fraîcheur, de nouveautés en mettant à disposition de l'ensemble des chercheurs des données communes de base pour leurs travaux, assorties de données spécifiques aux études particulières ; tout ceci référencé.

6.3 Un système SIG extensible aux autres domaines

Le système **extensible** est vu comme un SIG ouvert sur les autres domaines de recherche, pouvant être utilisé par l'ensemble des collaborateurs. Il regrouperait l'ensemble des données concernant les thèmes communs ou d'intérêt,. Il permettrait la consultation des données en commun par thème/section de recherche identique quel que soit le domaine. Exemple : étude du vignoble : Changins, Pullly, Gudo, Leytrons et Diolly, ainsi que des possibilités de liaisons extérieures comme par exemple aux données administratives, météorologiques, et **inter/intra domaine, inter/intra**

section de recherche. La partie « administrative » reste encore à discuter, elle ne fait pas partie du travail mais elle est tout de même proposée.

6.4 Conséquence : un léger changement dans l'organisation

Suite à la venue du SIG la mise en place d'une **stratégie d'organisation** s'impose. [*DESPONDS et al., 1999*].
Ainsi, chaque collaborateur devra « **rentrer ses données** » pour les parcelles étant placées sous sa responsabilité. Cette phase est redondante car ces données ont déjà été saisies par d'autres collaborateurs et ajoutées à leurs cahiers thématiques respectifs. Cependant, si chacun fait un peu, c'est tout le monde qui gagne ! et qui en retira les avantages ! l'utilisation du SIG sera alors possible.
De plus il faudra désigner un responsable, **un gestionnaire** qui entretiendra le SIG : réalisation d'une **maintenance** visant aussi bien au maintient des données mais aussi de leur structure.

6.5 Conséquence : un nouvel outil informatique

Le SIG sera un nouvel outil, informatique une fois de plus. Des considérations d'ordre technique apparaissent donc. Le SIG se devra d'être simple, souple, peu coûteux, voire avec une interface conviviale et efficace afin de ne pas commettre ni d'erreurs ni de perte de temps. . Les accès seront les mêmes que ceux actuellement en vigueur pour les données de la base Access VITIDATA, à savoir une **différentiation du mode de connexion selon l'usage**. En d'autres termes, un usage de consultation / traitement devra être différentié d'un usage d'insertion / modification de données.
Par la suite, l'idée est que chaque chercheur dispose du SIG à son bureau, qu'il utilise les blocs de données situées sur un espace disque commun, et sauvegarde personnellement ses propres données et traitements effectués. Ceci pour éviter la multiplication des fichiers de données sans connaissance de leur contenu, utilisation, génération…., la souplesse du SIG réside en ses aspects.
Il y a ainsi bon espoir de développement.

Résumé des points importants :

Un « super système » SIG n'est pas approprié pour la RVC. C'est un SIG additionnel, complémentaire et surtout modulable selon les essais et les modifications en cours d'expérience qui est proposé pour des données d'intérêt et communes.

Le SIG devra être extensible aux autres domaines de recherche.

Toutefois, le SIG apportera un nouvel outil qu'il faudra gérer par l'apport d'une nouvelle stratégie d'organisation.



Dans ce chapitre :

L'inventaire et la structuration des données,
Les acquisitions à réaliser et les données non nécessaires malgré leur existence quotidienne.
Des explications sur l'acquisition de la dimension spatiale des éléments du système d'expérimentation.

7 Analyse de l'existant, inventaire, structuration et acquisition des données.

7.1 Quelques considérations générales et démarche adoptée

La prise en compte de l'existant nécessite une bonne connaissance du fonctionnement actuel : ce qui a été fait § 4 : fonctionnement actuel de la RVC. Aussi, il est important que cet **inventaire** réponde aux besoins et attentes formulées par les collaborateurs ainsi qu'aux perspectives de fonctionnement du SIG, sa place au sein de la RVC : il ne faut pas arriver avec un projet tout cuit ! Ceci signifie que l'inventaire des données se décompose en deux parties. La première partie consiste a survoler l'ensemble des données (très nombreuses pour la RVC qui en possède aussi à la RAC…) puis dans une seconde partie de trier ces données afin de ne garder que celle qui sont utiles pour le SIG,

Rappel : les **objectifs de l'inventaire** sont les suivants: [*GOLAY, 2000*]

- 1 : identifier les informations utilisées sur les documents mentionnés par les utilisateurs, ainsi que le niveau de perception de ces informations,
- 2 : recenser les documents existants, leur qualité et leur contenu,
- 3 : permettre de dégager les goulots d'étranglement dans la circulation et la diffusion des informations.

Le DCA (**Diagramme conceptuel d'activité**) et le DFD (Diagramme de Flux de Données) situés § 4.7 servent encore une fois à mieux comprendre la situation particulière de la RVC.

Le choix du **niveau de détail** pour cet inventaire a été de se focaliser sur **des blocs de données** appartenant au même thème, et surtout faisant partie de la même base thématique au sein de la RVC. En d'autres termes, les spécifications du futur SIG ne sont pas abordées dans cette partie inventaire.

La démarche a été la suivante :

- 1 : **inventaire** des blocs de données,
- 2 : au sein même de ces blocs de données, **sélection, tri et structuration** des paramètres d'intérêt et/ou commun répondant aux besoins et attentes formulés, qui donneront plus tard les attributs. Le tri est réalisé directement car les chercheurs on l'habitude et voient immédiatement quelles données serviront oú non dans le SIG toutefois des conseils leur sont apportés. ,
- 3 : inventaire des **acquisitions** nécessaires.

Ainsi, il n'a pas été nécessaire de **restructurer les données** mais de bien les traiter par **blocs** (entité / ensemble cohérent) afin de trier l'utile de l'inutile pour le SIG en vertu des perspectives formulées. Enfin, des **acquisitions** sont prévues, notamment de cartes. La figure 10 ci-dessous récapitule le choix de démarche adopté.

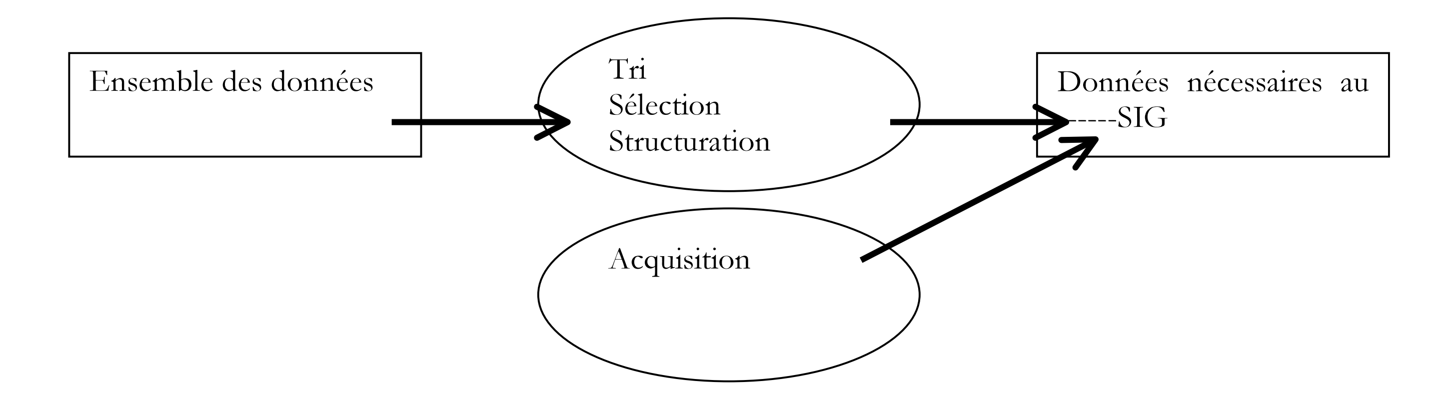


Figure 10 : le processus d'inventaire et de sélection des données, leur utilisation et les données produites.

7.2 Inventaire et structuration des données nécessaires

Ci-dessous table 10 sont décrits les **10 blocs de données** dans leur totalité, sans mention de leur contenu. Les blocs « externes » et « spécial » seront illustrés ici par les données météo et physiologiques respectivement Une simple description des blocs est fournie. (Le contenu apparaîtra dans le dictionnaire des données, lors de la modélisation du SIG).

BLOCS DE DONNEES	**FORMAT,** **remarque : toutes ces données sont disponibles en format papier, sous forme de cahier**	**ORIGINE / SOURCE**	**RESPONSABLE**	**DESCRIPTION DU CONTENU**
Données principales, communes et d'intérêt				
1 : DE BASE	Access	Base de données VITIDATA fournissant le cahier de vendange	RVC	Indication des paramètres principaux du système de conduite

2 : DE CONTROLE	Access, Word, Excel	Synthèse de diverses bases de données fournissant le cahier des résultats principaux des essais viticoles	RVC	Indication des paramètres de contrôle de végétation au cours de la croissance de la vigne
3 : DE VENDANGE	Access	Base de données VITIDATA fournissant le cahier de vendange	RVC	Indication des données de vendange
4 : DE VINIFICATION	Excel	Base de données de vinification fournissant le cahier des résultats d'analyse de vinification	RAC	Indication des données de vinification, mesurée au fil du temps de maturation / fermentation
5: DE DEGUSTATION	Word	Base de données des dégustations fournissant le cahier des résultats de dégustations	RAC	Indication des données de dégustation annuelle des vins
6 : EXTERNES exemple météorologique	Internet,	Rapport de météorologie et semblable	RAC	Indication des paramètres extérieurs, non relatifs à la conduite de la vigne mais ayant une importance. (pour ce travail : paramètres climatiques)
7 : PROJET SPECIAL Exemple de la physiologie	Excel	Base de données personnelles	Le chercheur responsable de projet	Indication des paramètres d'intérêt selon le projet, (pour nous la physiologie.)
8 : PROJET SPECIAL Exemple non disponible	Non disponible actuellement, mais envisagé			
Données spatiales				
PLANS RVC	Papier ou Excel	Base de données personnelles	RVC	Plans de situation du domaine et de l'intérieur des parcelles
CARTES DE BUREAUX D'ETUDE	Papier	Base de données personnelles	RVC	Cartes diverses en relation avec l'environnement du domaine
DE REPERAGE DE BASE	Papier	Cartes nationales et cadastrales	RVC	Utilisation en repérage

Table 10 : inventaire des blocs de données nécessaires au SIG

7.3 Les données non nécessaires au SIG et pourquoi.

Comme il y a été signalé au § 7.1, la méthode d'inventaire des données se passe en trois temps. Ici la table 11 illustre les données existantes, inventoriées lors de l'entretien, mais **non nécessaires** au futur SIG. Elles ne répondent pas aux perspectives futur du SIG. Une justification est apportée.

DONNEES NON-NECESSAIRES	POURQUOI ?
Données non spatiales	Les données non spatiales sont des données qui n'ont pas de relation avec le terrain. Comme cité § 6 les perspectives du futur SIG sont axées sur un moyen additionnel orienté « terrain » d'accès aux données et fonctions spatiales. (Exemple listing de publications)
Données or RVC	Le but du SIG n'est pas de s'occuper de l'ensemble des données de la RAC liées à la viticulture, mais uniquement des données communes et d'intérêt utiles à la RVC. (Exemple : les données des essais de la station de Changin).
Bilans interannuels	Les bilans interannuels constituent des données de synthèse sur plusieurs années, il est donc possible de retrouver ses données, brutes, en consultant différents cahiers de différentes années. Ces bilans sont établis selon les thèmes et les besoins, ils sont sous forme de rapport thématiques. (Exemple : bilan interannuel des essais physiologiques)
Les publications	Les documents publiés (de sortie) constituent à leur tour des données. En effet leur contenu est une synthèse thématique de multitudes d'essais. Le SIG ne peut pas comporter ce genre d'informations pour la simple raison que se sont des rapports sur les données. (Exemples : rapport d'activité, article de magazine…)
Les projets temporaires particuliers	La multitude des projets ne justifie pas une centralisation absolue de toutes les données produites, surtout de celles qui ne sont utilisées qu'occasionnellement. (Exemple :Etude de l'échange gazeux des feuilles)
Les interventions effectuées	Par interventions on entend travaux effectués. Ils sont identiques et réguliers sur l'ensemble du domaine, ils sont aussi classiques pour la vigne. On rappelle que la RVC n'étudie pas l'état phyto-sanitaire de la vigne (contrairement à la RAC). Ainsi ils ne sont pas utiles car pas étudiés. (Exemple : traitement fongique)

Table 11 : inventaire des données existantes mais non nécessaires au SIG

7.4 Acquisitions dont la dimension spatiale

Après avoir réalisé l'inventaire des données, il faut faire l'inventaire des **acquisitions à prévoir**. Ces acquisitions doivent répondre aux perspectives du futur SIG, besoins et attentes des chercheurs. Ce sont :

- Le parcellaire du domaine du Caudoz, Cf. fig. 3 du début, forme papier
- Les plans de la structure interne des parcelles, autrement dit les plans de plantation des vignes, Cf. Fig. 5 du début, forme papier

- Les cartes relatives à la représentation spatiale d'un phénomène temporel, carte des sols, des profils et sondages, des teneurs en eau du sol, et de profondeur de sol, sous forme informatique afin de ne pas altérer la précison,
- Une orthophoto du domaine pour permettre la digitalisation en complément au plan du domaine. Cf. fig. 11 ci dessous.

Afin de mettre en œuvre le SIG, la **dimension spatiale** doit être acquise. Selon quelle **précision** ? En vertu des besoins et des objectifs des chercheurs, la volonté d'avoir une mise en situation spatiale est dominante. De plus, pour des raisons économiques, un relevé des parcelles est à proscrire. Ainsi, une **digitalisation des parcelles selon l'orthophoto** locale semble être la méthode la plus simple, la plus économique et surtout la plus fidèle à la réalité. Des retouches locales pourront avoir lieu, et alors la **précision d'environ 3 mètres** sera atteinte, ce qui suffit amplement. En effet, on pourrait même penser à une structure spatiale purement topologique (c'est à dire un simple agencement de carrés représentant des parcelles sur un domaine délimité). Cependant, si l'information géographique peut être atteinte facilement, pourquoi s'en priver ? L'utilisation du SIG n'en sera alors que plus fidèle, ce qui facilite beaucoup de choses, et nous restons ainsi en **adéquation entre le SIG et le terrain**. Pour ce qui est des lignes de vigne, elles sont acquises selon les plans parcellaires, il en est de même pour les variantes, les répétitions et les pieds. La position exacte des pieds ne serait pas justifié, aussi bien économiquement que pour les besoins du système !. La structure pour ces éléments est donc plutôt **topologique** mais que partiellement, car ils s'inscrivent dans une parcelle qui, elle, est référencée. (Le § 9.4.1 explique l'acquisition, par digitalisation, par rapport à l'orthophoto et le cadastre).

Résumé des points importants :

les données étant très nombreuses et diversifiées, elles ont été structurées en 10 blocs de données appartenant au même thème, thème existant déjà au sein de la RVC afin de ne pas modifier les mécanismes mis en place au niveau de la gestion des données.

Les acquisitions de données sont peu nombreuses et rapides à obtenir, la dimension spatiale est acquise grâce à une orthophoto régionale. Les éléments du système d'expérimentation sont acquis sur la base des plans de parcelle, sous forme semi topologique. Le tout forme un ensemble harmonieux et fidèle de la réalité.





Figure 11 : orthophoto du domaine de la RVC.

Dans ce chapitre :
Des explications concernant la modélisation de la problématique.
Les Modèles Conceptuels et logiques de données.
Les éléments créés pour le SIG ainsi que les nouveautés apportées.

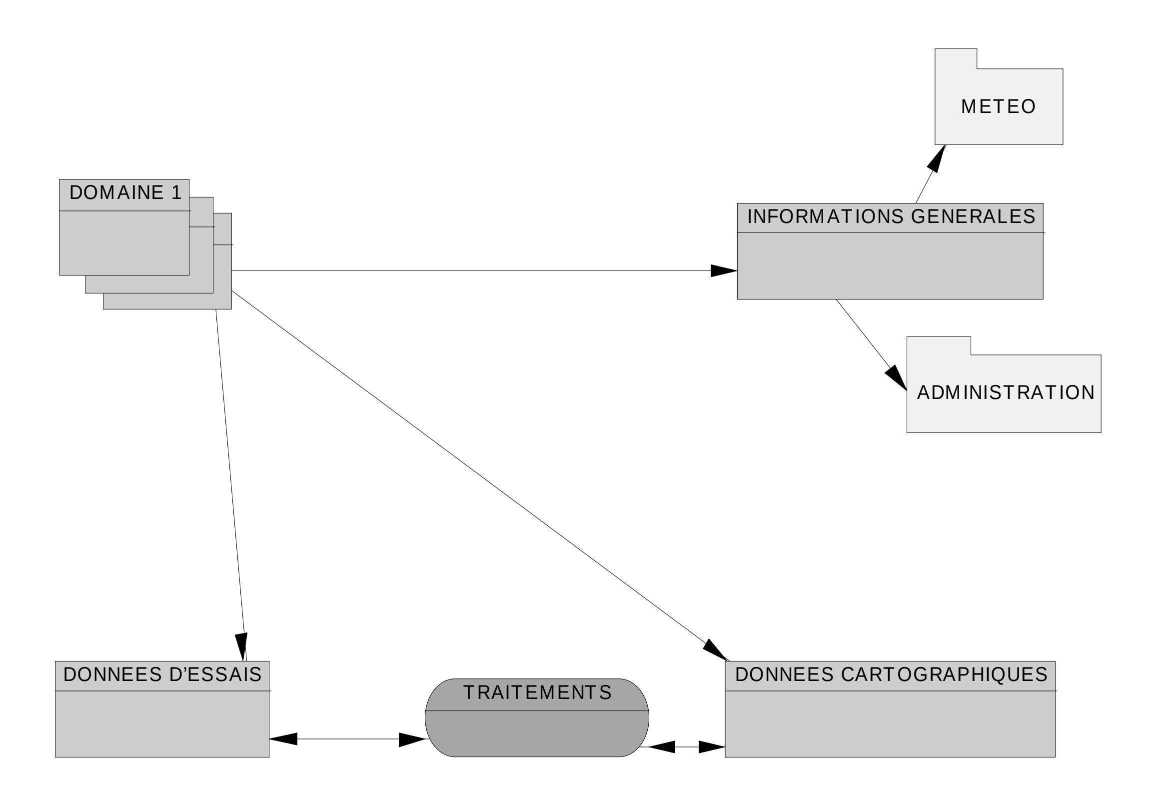
8 Modélisation / conception du SIG

8.1 Rôle du SIG ?

Avant de réaliser la modélisation de la problématique et de l'implanter sur logiciel SIG, il est nécessaire de connaître les différents **rôles** du SIG. Ces rôles du SIG répondent aux besoins et attentes des chercheurs. Les figures 12 et 13 ci dessous illustrent les rôles du SIG au niveau de l'ensemble des données.

Au niveau des groupes des données (fig. 12), il est nécessaire de ne pas mélanger les thématiques. Ainsi, il est préférable d'isoler les données d'essais des données générales et des données cartographiques. Trois **groupes de données** de même thématique sont donc créés. Le quatrième groupe est celui concernant le « domaine ». Ce dernier donne accès à l'ensemble des domaines de recherche, ce qui fait du SIG un SIG modulable et extensible à l'ensemble des stations fonctionnant sur le même principe. C'est à partir du domaine que les autres groupes de données sont atteignables. Les noms donnés à ces groupes de données sont : « le domaine », « données cartographiques », « infrastructures et météorologie », et enfin « stations de recherche / essais ». Entre ces groupes de données, reliées à l'entité centrale « domaine », des traitements pourront avoir lieu entre les données d'essais et les données cartographiques. Ces traitements sont le plus souvent des calculs statistiques réalisés avec le logiciel WIDAS ou des interprétations purement scientifiques Aussi, la jonction à des bases de données extérieures doit être prévue, notamment pour les données météorologiques et administratives, mais n'est pas envisagée pour le présent travail de diplôme.

ROLES DU SIG AU NIVEAU DES GROUPES DE DONNEES: UNE IMPLICATION POUR LA MODELISATION



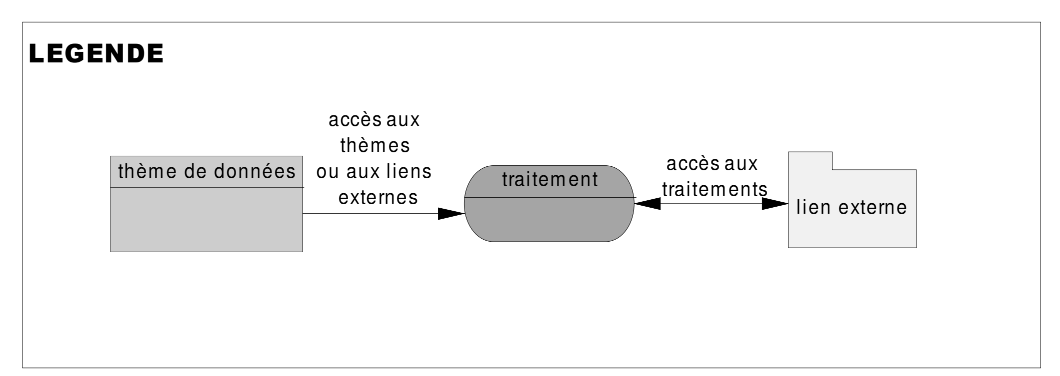


Figure 12 : les rôles du SIG au niveau de groupes de données.

Pour ce qui est des blocs de données (Cf.§7.2) et (table 10). Il est important de pouvoir distinguer les fonctions des différents blocs. Ces fonctions sont :

- **Fonction d'objet géographique** : afin de dissocier l'objet géographique de l'objet d'essais pour permettre la mobilité spatio-temporelle des objets,
- **Fonction de données de données de base** : afin d'informer sur les données dites de base fixes durant plusieurs années pour le même objet d'essais,
- **Fonction de données d'essais** : afin d'informer sur les données annuelles correspondant aux relevés des essais.

La figure 13 classe les tables de données en effectuant cette séparation fonctionnelle. On y distingue les cinq entités du système d'expérimentation (parcelle, répétition, variante, ligne , pied) classés avec leurs trois fonctions énumérées ci-dessus. L'ensemble des tables est **arrangé selon cette matrice** ainsi formée. Notons que :

- Il y a cinq **tables objets** référencées correspondant au système d'expérimentation, qui renseignent sur l'objet géographique lui-même.
- Leur correspondent cinq **tables tabulaires de données de base pluriannuelles** correspondant aux données de base liées à l'objet d'essais et non plus à l'objet géographique.
- Pour chaque entité du système d'expérimentation, plusieurs **tables tabulaires comportant les données d'essais thématiques annuelles** viennent se greffer à partir des données de base afin de livrer les données annuelles d'essais.

Les éléments suivants **influencent aussi la conception** et le contenu de la base de données :

- Les requêtes que les chercheurs veulent réaliser,
- La particularité de permettre de réaliser des adaptations d'essais en cours d'expérimentation,
- Le problème spatio-temporel à prendre en compte,
- Le respect de la structure proposée,
- Le respect du contenu proposé,
- Les autres volontés particulières des chercheurs,
- L'affichage des essais en cours d'exploitation,
- L'archivage des essais passés
- etc.…

ROLES DU SIG AU NIVEAU DES TABLES DE DONNEES:
UNE IMPLICATION POUR LA MODELISATION

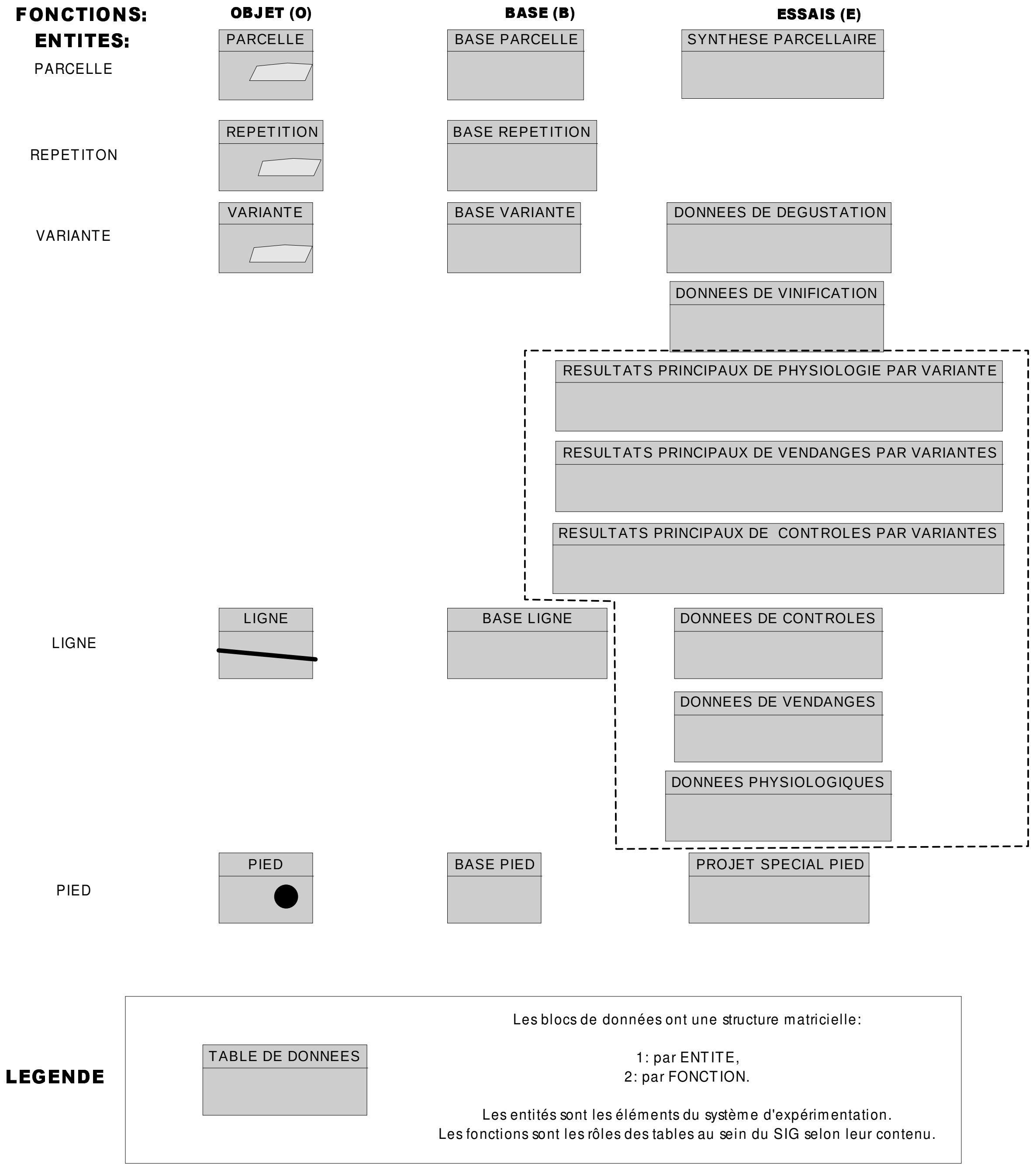


Figure 13 : les rôles du SIG au niveau des tables de données.

Ces **pré-requis** doivent être pris en compte avant de modéliser car ils impliquent la modélisation. Celle-ci peut alors commencer.

8.2 La modélisation conceptuelle : le MCD

Une fois ces différentes étapes franchies (besoins, attentes, rôles, perspectives,…,) il est nécessaire de passer à la **transcription des informations** relevées au niveau de la réalité aux informations nécessaires. Cette étape est la modélisation *[GOLAY, 1999* & *2000], [PLAZANET, 2000], [SPACCAPIETRA,2000]*.. Elle consiste à une représentation simplifiée de la réalité, dépendant du concepteur et de sa perception, afin de pouvoir passer ensuite à une **implantation** de cette réalité modélisée sur un logiciel SIG. Pour parvenir à cela on utilise un MCD : un Modèle Conceptuel de Données. Celui ci à deux principales finalités :

- Représenter synthétiquement et simplement le résultat de la recherche effectuée (analyse des besoins et des attentes, perspectives de développement et les rôles du SIG…) afin de le monter aux futurs utilisateurs et de discuter avec eux, pour détecter d'éventuelles discordances et d'éventuels problèmes,

- Modéliser le résultat de travail de recherche de façon complètement indépendante d'un logiciel SIG, ce qui permettra de choisir de manière optimale les différents supports informatiques à utiliser pour l'implantation du SIG.

Le MCD utilise un **formalisme**, c'est à dire une façon connue, codée de représentation des données. Le formalisme choisi est le formalisme **« entité-relation ».** Il est composé des éléments suivants :

- Des entités

- Des relations

- Des attributs dont des identifiants

- Des cardinalités

Le modèle « entité _ relation » est expliqué à la table 12 ci dessous :

Les données sont structurées sous forme de tables / entités (rectangles) reliées entre elles par des relations. Le nom des tables figure dans la partie haute de chaque rectangle. Les attributs sont listés dans le corps du rectangle : ils représentent les informations se rapportant à la table. Les attributs soulignés sont des identifiants, c'est à dire les attributs permettant de « retrouver » un objet de la table grâce à son unicité devant être garantie. Chaque table a une référence géographique (illustrée par un icône dans le coin droit des entités, point pour ponctuel, ligne pour linéaire et surface pour surfacique) correspondant à une couche dans le SIG, c'est à dire à un ensemble spatial d'objets géographiques appartenant à une même thématique. *[PYTHOUD, CALOZ, 2001]*.

Table 12 : description du modèle entité / relation

Le MCD est décrit § 8.3. Il a été réalisé grâce au logiciel « outil case » **POWER AMC 7 (**et mis au propre grâce à EXCEL) permettant de représenter correctement les différents éléments composants le MCD.
Le MCD figure en annexe 5.

8.3 Description du MCD

8.3.1 But du MCD

Comme cité plus haut le MCD permet la représentation simplifiée de la réalité perçue, demandée, nécessaire…afin de pouvoir l'implanter dans le SIG. Ici, quelques éléments importants doivent être pris en considération, ce sont :

- **L'appréciation du degré de détail** : seul les attributs les plus pertinents ont été conservés. Ce sont les attributs communs et d'intérêt.

- **Le respect des groupes et des blocs de données** : selon les thématiques abordées, les données sont structurées afin de ne pas se perdre, et que les recherches soient intuitives, c'est à dire sans avoir à lire et comprendre une structuration complexe.

- **La possibilité d'extension** : l'ajustement des tables, et leurs relations permet l'extension aux autres domaines de recherche, ceci par le biais de l'entité centrale rattachant toutes les autres : « domaine ». Il en est de même pour l'accès au serveur météo et aux données administratives stockées sur d'autres bases.

- **La possibilité d'être modulable** : les tables supplémentaires peuvent être crées, dont deux ici sont déjà prêtes afin de recevoir des attributs d'éventuels projets ou de cartes nouvelles.

- **La volonté de rester proche de l'état actuel de fonctionnement** : cette idée déjà développée §6.1 et §6.2 permet au SIG de ne pas être quelque chose de tout neuf, mais plutôt de complémentaire sur une base connue.

- **La simplicité de retrouver les éléments**, d'une part grâce à la partition du territoire qui crée l'individualité d'un objet lorsque celui ci est inclus dans un autre, et d'autre part par des identifiants qui respectent la codification actuelle. Ainsi, il n'est pas introduit d'identifiant « tout nouveau » afin de ne pas troubler la réflexion qui se veut simple et déductive, allant du général au particulier. Un exemple : trouver la ligne numéro 89764 n'aurait pas eu de sens, mais trouver la ligne 4 de la variante G, de la parcelle 6 sur le domaine de Caudoz, n'est pas difficile, on peut la trouver soit même !

Les § ci-dessous donnent une brève description des entités et des relations du MCD, le **dictionnaire des données** étant toutefois la référence.

8.3.2 Les entités

- L'entité **domaine** est l'entité centrale du MCD. C'est à partir d'elle que toutes les autres entités sont reliées, ceci pour la simple raison que le SIG se doit d'être modulable et extensible aux autres stations de recherche. L'entité comporte quelques informations brèves, son identifiant lui est propre, le nom du domaine et son abréviation apparaissent aussi . Son rôle est central, elle est le cœur de la base de données.

- **Les entités du groupe de données « infrastructures et météorologie »,** représentent les objets géographiques présents sur le domaine : bâtiments, station météo, annexes, laboratoires, parcelle libre, routes, les critères zonaux administratifs et enfin les points d'études divers tels que sondages et profils, qui contribuent à l'élaboration des cartes, sans en faire partie directement. Ici, la

représentation complète des objets constituants le domaine a été choisie afin d'être en accord avec la réalité mais aussi et surtout de pouvoir accéder aux données. Ainsi, les données météorologiques sont accessibles par l'objet « station météo », les données administratives par les « bâtiments », l'accès au domaine et l'usage des routes par « routes », etc.…. Leur rôle est de permettre un accord complet avec la réalité et de faciliter l'accès aux données de manière simple et intuitive même si l'on ne connaît pas le domaine en question, ce qui est le cas assez régulièrement. Les données mentionnées sont entre autre : nom, abréviation courante, le responsable, l'usage, .etc…

- **L'entité du groupe « données cartographiques »** » est unique : « unité de sol ». Elle permet l'apport complet des données pédologiques, en mentionnant les caractéristiques indispensables telles que nom de sol, profondeur, hydromorphie, roche mère…. Son rôle est principalement triple : à la fois de permettre l'accès aux informations cartographiques, d'autre part de permettre l'édition de cartes illustratives et de cartes thématiques, et pour finir d'être le support de base de traitements et d'analyses spatiales. L'unité de sol représente le plus petit ensemble zonal ayant des caractéristiques identiques sur un ensemble de paramètres définis. Autrement dit, il suffit d'un seul paramètre différent pour conduire à la génération d'une autre unité. Ces unités de sol ont été définies par une entreprise spécialisée en pédologie. La table se compose de codes pour chaque paramètre, suivis de leur signification, afin de ne pas avoir recours à la décodification en tout temps, mais aussi de permettre d'inscrire la valeur effective du paramètre et non sa classe d'appartenance. L'annexe 3 donne la décodification des unités de sol. Reste à noter que ces unités son identifiées par une étiquette se basant sur les paramètres essentiels décrivant objectivement d'unité. Un identifiant unique est joint à chaque unité. (ID).

- Enfin, la plus grande partie des données fait partie **du groupe « station de recherche / données d'essais ».** Les cinq entités composant ce groupe sont les éléments du système d'expérimentation décrit § 4.6 . Ce sont : la parcelle, la variante, la répétition, la ligne et enfin le pied. Ils représentent ici une partition du territoire, c'est à dire un découpage successif du territoire sans recoupement possible. Chaque entité contient des blocs de données, c'est à dire un ensemble de données issues d'une même source, cf.§7.2.. Maintenant, il est important de bien comprendre la chose suivante : afin de distinguer les rôles des tables selon leur fonction (objet, base, et essais) (Cf. § 8.1), chaque table géographique du système d'expérimentation est liée respectivement à sa table des données de base qui est elle-même liée à une ou plusieurs tables thématiques des données d'essais. Ceci donne alors une structure à trois niveaux (objet, base, essais) coordonnée dans l'espace et dans le temps. En effet, les objets géographiques sont dissociés de leurs attributs pluriannuels eux-mêmes distingués de leurs attributs annuels. Ceci permet une acceptation des modifications d'essais en cours d'expérimentation et permet aussi la prise en compte du problème spatio-temporel. Les données formulées sur l'entité « pied » restent sur le cep : elles permettent ainsi l'étude au cas par cas. Il n'y a pas de synthèse des informations des ceps à la ligne. Les lignes comportent les données qui y sont relatives, il en est de même pour les répétitions et les variantes. Ceci est évident, mais le chercheur a la possibilité de saisir les informations au niveau qu'il le souhaite, ce qui lui laisse libre choix. Ainsi, par exemple, il peut saisir les données ligne par ligne ou bien directement par variante, ce qui est beaucoup plus rapide et qui correspond mieux a la situation actuelle. Ainsi, les tables entourées de pointés de la figure 13, sont regroupées, car elles contiennent les mêmes attributs Toutefois, on ne retrouve la structure identique de la base de données VITIDATA que si « il n'y a qu'une seule ligne par variante pour chaque répétition ». Le lecteur réfléchira sur cette phrase : elle permet de saisir correctement l'esprit de conception, et les chercheurs l'esprit de classement : ce qu'ils appellent communément « une répétition » devient ici « une variante dans une répétition » Ensuite, des variantes aux parcelles, une interprétation libre est possible par les chercheurs. Elle permet de faire un bilan, par essai, donc par parcelle, sur les différentes variantes composant l'essai. La répétition ne comporte pas beaucoup d'informations, car on rappelle que son unique rôle est de

permettre le « brassage »statistique des données afin d'obtenir un échantillon représentatif, pouvant d'ailleurs être moyenné sous cette condition. Le rôle fondamental de ce groupe de données est de permettre la récolte des informations aux différents niveaux du système d'expérimentation, de les synthétiser automatiquement là où on le souhaite, et enfin d'élaborer un bilan sommaire concluant les variantes des essais au niveau de la parcelle, mais aussi des bilans interannuels en un clin d'œil !

8.3.3 Les relations

Les relations entre les entités sont pour la plupart partantes de l'entité centrale « domaine ». Ceci permet de se rattacher à un élément semblable pour toutes les entités. On trouve quand même une structure circulaire entre les entités du système d'expérimentation à partir de l'entité « parcelle ». Toutes **les cardinalités sont de type 1 :1** car **la partition du territoire** domine pour chaque relation. Ainsi, on obtient des cardinalités de ce type pour chaque relation. Ceci aura un impact lors de la traduction du MCD en MLD § suivant. La simplicité s'est voulue être prépondérante de manière à comprendre le MCD et donc la structure de la base de données, ce qui la rend accessible. Ceci se visualise clairement par la liaison entre les tables « objets géographiques » vers les tables « bases » puis enfin vers plusieurs tables « essais ». Ceci forme une chaîne à partir de chaque entité du système d'expérimentation. La figure 14 ci-dessous l'illustre :

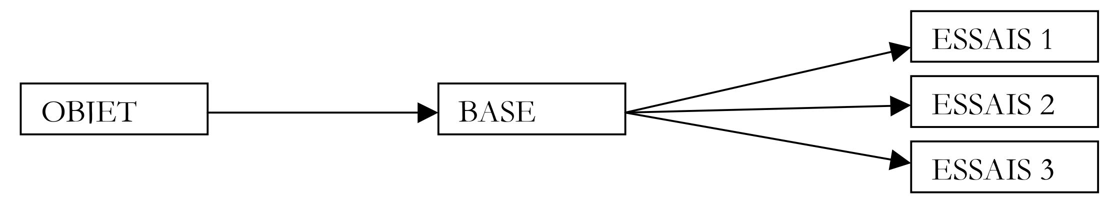


Figure 14 : la chaîne entre tables du système d'expérimentation : Objet >> Base >> Essais

8.4 Traduction de MCD en MLD[16]

Afin de permettre l'implantation de la base de données dans un logiciel, il est nécessaire de traduire le MCD en un autre modèle. En effet, comment implanter le MCD ? : il faut le **traduire** en une forme permettant sa compréhension afin de pouvoir **l'implanter physiquement** par la suite dans le logiciel choisi. Nous ne développons pas ici les règles qui régissent cette traduction. Toutefois, il est important de signaler que les cardinalités 1 :1 présentes au niveau de chaque relation du MCD, permettent de ne pas créer de tables supplémentaires issues de relations à autres cardinalités. Ainsi, les tables sont celles des entités, sans tables de relation intermédiaires. Les identifiants constituent un moyen unique de repérage. Ils s'ajoutent aux tables entités afin de permettre le suivi des relations sans que des tables relations soient crées. Le MLD figure annexe 6.

8.5 Les identifiants

Ce paragraphe a pour but **d'expliquer les identifiants choisis pour chaque table de même fonction et pourquoi.**.

- Les **tables OBJETS** font toutes intervenir un ID, c'est à dire un identifiant unique pour toute la base de données. Cette ID est le résultat d'un numéro automatique.

[16] MLD : Modèle Logique des Données

- Les **tables BASES** comportent un *code_nom de l'objet* auquel se rattache les données de base, sous la forme txxxrxxx, x et y étant des chiffres. Ceci signifie « table xxx record yyy », unique pour toute la base de données. On peut ainsi multiplier les données de base pour le même objet, ceci sur plusieurs années. Un exemple : l'objet est le même, l'essai change : les données changent toutes elles-aussi. ; il en est de même si les conditions d'expérimentation évoluent en cours d'essais : un nouveau code est alors de rigueur faisant appel à des années différentes. Il y donc un code unique à chaque fois qu'il y a au moins un attribut de la table qui change.

- Les **tables ESSAIS** comportent un *ID_nom de la table d'essais*. Les mêmes considérations que précédemment sont valables. Toutefois, ici cet identifiant n'est pas repris par liaison à d'autres tables. Ainsi, sa forme est *code_nom de l'objet-année d'essai* . il permet d'identifier une ligne (record) et une seule pour toute la base de données. la recherche est donc plus facile et plus rapide que par succession d'identifiants.

- Des **LIENS LOGIQUES** permettant de se repérer rapidement dans les essais, sont créés par liaison à la table « base parcelle ». Ce sont le nom du domaine, le numéro de la parcelle d'essai, et le titre de l'essai.

8.6 Les nouveautés

Les nouveautés apportées par le SIG sont les suivantes : elles sont, soit totalement nouvelles, c'est à dire inexistantes précédemment, soit vues d'une autre façon, ce qui crée aussi une nouveauté par rapport à l'état de fonctionnement actuel de la station de recherche.

- Il y a une distinction entre l'objet et ses données laissant une grande souplesse d'exploitation,
- La récolte des informations se réalise par lignes puis par variantes, et non plus par répétition comme dans la base de données VITIDATA,
- Le choix de saisir les données selon le niveau du système d'expérimentation (pied, ligne, variante, parcelle) est possible,
- La possibilité de saisir les données à plusieurs niveaux à la fois est possible : ainsi les lignes appartenant à la même variante constituent soit des données ligne par ligne soit une synthèse (des lignes) au niveau de la variante les contenant,
- Des interprétations ont lieu pour les variantes afin de pouvoir les différentier, ces interprétations sont libres, elles doivent cependant permettrent le repérage des éléments discriminants,
- Des bilans et conclusions d'essais sont accessibles au niveau des parcelles,
- L'accès aux données d'intérêt et / ou communes est immédiat, sans se perdre dans les données secondaires,
- L'apport de nouveaux attributs jusque là sous-entendus ou inexistants,
- La possibilité de stocker des informations, non seulement à chaque niveau du système d'expérimentation, mais aussi pour n'importe quelle durée, allant du groupe d'années au jour comme pour les données physiologiques,

- Il est possible de consulter les données sur plusieurs années, les bilans interannuels sont visibles dans la même table et non dans plusieurs cahiers séparés,
- Des attributs d'interprétation permettent de discriminer les variantes entre elles, et ceci même sur plusieurs années,
- Les requêtes de sélection permettent de trier les données afin de choisir celles qui intéressent,
- L'ensemble des données est centralisé,
- L'accès aux autres domaines est possible,
- Enfin, la visualisation spatiale apporte la dimension manquante, et permet toutes les manipulations induites.
- Etc…

8.7 La création d'indices

Les blocs de données de dégustation et de vinification, représentent une finalité des essais dans le sens que leurs résultats sont discriminants. En effet, la dégustation met en évidence les vins significativement différents au niveau de certains qualificatifs et la vinification peut détecter une vinification normale d'une vinification problématique. Il n'en est pas de même pour les autres blocs de données qui représentent des données plutôt brutes, nécessitant une synthèse et une interprétation.

8.7.1 Indice de dégustation, Id

Pour les données de dégustation, il a été opportun de créer un **indice qualifiant la variante**. Ceci pour un but : obtenir un indice permettant de **distinguer les variantes entre elles**, et de mettre ainsi en évidence de potentielles différences significatives. Ceci est une **discrimination.**

Ceci a été réalisé pour les données de dégustation, au niveau des groupes de Duncan. Les vins issus d'un même essai sont dégustés et classés en groupes par une à plusieurs lettres, appelés groupes de Duncan. Un vin est dit significativement différent d'un autre lorsque aucune de ses lettres le composant ne se retrouve dans celles d'un autre vin.

Pour les données de dégustation, le développement d'un indice a été réalisé et les explications figurent ci-dessous. On le nomme Id. l'échelle est cardinale, l'unité en [%].

La table 13 : relate l'explication de l'indice de dégustation Id.

Id [%] se calcule comme : (Au niveau des groupes de vins, nombre de fois que la variante est significativement différente de ses collègues / nombre de fois que cette variante peut être significativement différente au maximum)*100. On peut dire ceci de la manière plus simple pour une variante : (nombre de fois qu'elle est différente / nombre de fois qu'elle peut au maximum l'être)*100. Un vin est considéré significativement différent des autres lorsque les groupes de vins établis lors de la dégustation ne contiennent pas deux fois la même lettre. Ceci est réalisé pour chaque qualificatif faisant apparaître une différence significative entre les notes

attribuées par les dégustateurs. Ainsi, le nombre de fois qu'une variante peut être significativement différente de ses voisines vaut : le nombre de qualificatifs significativement différents * (nombre de variantes-1). Le terme –1, s'explique par le fait qu'une variante ne peut pas être différente d'elle-même, il faut donc la retrancher.

En adoptant les notations suivantes on trouve une formule simple pour l'indice.

Soit V le nombre de Variantes composant l'essai,

Soit QSD le nombre de Qualificatif Significativement Différent,

Soit S la Somme des différences significatives de la variante comparée à ses collègues, représentant le numérateur,

Soit N = QSD*(V-1), représentant le dénominateur

Alors Id= (S/N)*100, [%], remarque: si N=0 alors Id=0

Si Id =0 alors la variante n'est pas significativement différente de ses collègues, au contraire si Id=100, la variante est totalement différente de ses collègues (elle a cumulé toutes les différences possibles sur tous les qualificatifs !).

Table 13 : description du mode de calcul de l'indice de dégustation. (Id)

Le tableau 14 ci-dessous illustre par un exemple concret les calculs pour la page 85, du cahier des dégustations de 2001, qui fait apparaître des différences significatives pour trois qualificatifs, l'essai comportant 4 variantes :

Numéro variante	Code variante	Groupe vin selon le qualificatif « évolution »	Groupe vin selon le qualificatif « fruité »	Groupe vin selon le qualificatif « finesse »		
1	01.353	BC	A	A		
2	01.354	C	A	AB		
3	01.355	AB	A	B		
4	01.356	A	B	C		
Numéro variante	Code variante	Id selon le qualificatif « évolution »	Id selon le qualificatif « fruité »	Id selon le qualificatif « finesse »	Bilan variante	Id
1	01.353	1/9	1/9	2/9	4/9	44%
2	01.354	2/9	1/9	1/9	4/9	44%
3	01.355	1/9	1/9	2/9	4/9	44%
4	01.356	2/9	3/9	3/9	8/9	88%

Avec: 9= 3(4-1) >> Il ressort que la variante 01.356 est plus différente de ses collègues : elle se démarque*

Table 14 : exemple de calcul de l'indice de dégustation Id.

8.7.2 Indice de vinification

Contrairement à l'indice de dégustation, l'indice de vinification est basé sur une échelle ordinale. Il se rattache à la variante et est associé à ses collègues du même essai au sein de la synthèse parcellaire d'essai.. Les chercheurs établissent un qualificatif binaire permettant de déterminer la qualité, le déroulement et l'état de la vinification. Les termes proposés sont :

- Vinification NORMALE, signifiant « en ordre »

- Vinification PROBLEMATIQUE, signifiant qu'il reste du sucre ou bien qu'un goût réduit est présent…

8.8 Le dictionnaire

Le dictionnaire des données est l'élément indispensable du SIG. Il permet de **regrouper l'ensemble des définitions des attributs et tables constituant les couches du SIG**. Il fait aussi apparaître les méta données, c'est à dire les données sur les données présentes en tant qu'attributs dans le SIG. Le dictionnaire annexé est complet. Il représente **une référence de base pour le SIG**. Les utilisateurs devront le conserver à leurs côtés afin d'utiliser correctement le SIG, de comprendre sa structure et son mode de fonctionnement. On peut dire aussi que la visualisation du MCD aide la compréhension sans apporter les éléments introduits par le dictionnaire. Le dictionnaire figure en annexe 1.

Les méta données surchargent la base et créent une redondance, elles ne sont pas directement utilisées mais plutôt d'intérêt second particulier. Leur connaissance est indispensable et sauve bien des situations ! (Exemple : la méthode utilisée pour la récolte de données physiologiques ne figure pas dans les données directement (ce n'est pas un attribut) mais elle devrait figurer dans les méta données afin de comprendre ce qu'a fait la personne en charge de la saisie qui n'est pas forcément celle qui analysera les données !)…Il faut signaler les données sur les données, mais ne pas rentrer pour autant dans les détails.

Résumé des points importants :

L'ensemble des étapes de la modélisation a été réalisée de manière à laisser une grande souplesse et adaptation du SIG afin de répondre à la plus grande partie des demandes tout en restant assez simple et intuitif.



Dans ce chapitre :
Après le passage de la réalité à la modélisation au chapitre 8 précédent, voici le passage de la modélisation à l'implantation sur SIG : le prototypage.
Cette étape décrit le système exploratoire mis en place qu'est le prototype.

9 Réalisation du prototype SIG

9.1 Choix du logiciel

Le logiciel choisi est le logiciel **SIG MANIFOLD**. Ce logiciel a été choisi pour plusieurs raisons :

- La licence d'exploitation a un prix modeste d'environ 250.-US$, ce qui est abordable pour une organisation de recherche comme la RVC pour équiper au moins un poste informatique,
- Le logiciel dispose d'une partie « dessin » et d'une autre partie « projet » sous forme d'explorateur Windows, permettant ainsi de connaître en tout temps la hiérarchisation des données au sein du projet,
- La société propose de nombreux services et extensions qui peuvent s'avérer utiles par la suite,
- De plus la société semble être assez proche de ses utilisateurs, ce qui revient à une idée de « service après vente » accessible, ce qui est toujours agréable,
- Tout le projet tient en un seul fichier (extension .map), inutile de rechercher les couches et de les ouvrir à chaque session, tout est déjà structuré et intégré au fichier.

9.2 Méthode d'importation des données dans MANIFOLD

9.2.1 Projection et système de coordonnées

Afin d'être compatible avec les autres stations de recherche et les autres domaines sur toute la Suisse, le système de **projection Suisse** a été choisi. Pour ce qui est des coordonnées, leur système est le **système national** réglementaire **(**avec Bern ayant pour coordonnées (Y=600'000 ;X=200'000)). Le tout est donc dans le système **CH1903**, utilisé par la mensuration officielle.

9.2.2 Importation de l'orthophoto

L'orthophoto servant à la digitalisation des parcelles d'essais (Cf. § 7.4 et 9.4), elle doit être importée dans le système CH1903. Le logiciel MANIFOLD permet les transformations : il faut lui stipuler que notre orthophoto de la zone du Caudoz doit être en **New Swiss Grid**.

9.2.3 Importation des données cartographiques

Les données cartographiques importées directement (donc non digitalisées) sont les données de la couche « unité de sol ». Cette couche a été élaborée par une entreprise de pédologie grâce au logiciel ARCVIEW. Son importation ne pose par de problème de reconnaissance de fichier, mais on doit spécifier « New Swiss Grid » afin de rester dans le système CH 1903, et d'être ainsi compatible avec toutes les autres couches.

9.3 Structure du projet et des tables dans MANIFOLD

9.3.1 Hiérarchisation du projet

Les dossiers (folders) respectent la structure du MCD. Ceci facilite l'exploration des données. Il y a adéquation entre le MCD et la structure au sein du SIG dans la mesure du possible (seuls quelques noms différent ou bien des couches sont structurées d'une façon légèrement différente du MCD, pour des raisons de prototypage). .. L'annexe 7 présente la totalité de la hiérarchisation du projet. On distingue les :

- **Dossiers**, qui représentent les groupes de données,
- **Drawing**, qui sont en fait les objets référencés,
- **Tables associées au drawing**, qui contiennent les attributs des objets,
- **Tables** non associées, ou libres, qui contiennent les blocs de données, (Fig.13 et 14) ,
- **Scripts** qui sont des fonctions réalisant des opérations sur les tables (Cf. § 9.7)
- **Cartes**, qui sont des regroupements de plusieurs drawing dans une même vue,
- **Query**, qui sont des requêtes préparées pour effectuer des sélections prédéfinies,
- **Layout,** qui sont des masques d'impressions prédéfinis.

9.3.2 Structure des tables : quelques éléments particuliers.

Comme expliqué § 8.1 et 8.3.3 ainsi qu'aux figures 13 et 14, les tables se divisent en tables OBJET, tables BASE, et tables ESSAIS. Selon chaque division, des caractéristiques leurs sont propres : voir annexe 8.

- Pour les **tables OBJET** : un ID unique pour toute la base est attribué à chaque objet géographique, les ID des autres objets liés sont ajoutés comme le veut le modèle logique des données (annexe 6) , un attribut nommé « actif » pouvant comporter les valeurs binaires [oui, non] permet de gérer l'affichage à l'écran en ne montrant que les objets oú cet attribut vaut [oui], les attributs « année début objet » et « année fin objet » permettent de connaître la durée de vie de l'objet et par la même occasion de réaliser des requêtes afin de retrouver des objets ayant existés par le passé, (un nouvel objet étant crée après une période d'interruption), enfin le numéro courant utilisé par les chercheurs est joint à partir de la table de base correspondante. **Ces quelques attributs, de chaque table OBJET, permettent de retrouver des objets, de visualiser les objets courants et de réaliser un historique spatio-temporel.**

- Pour **les tables BASE** : on trouve les attributs valables pluri annuellement concernant les données de base relatives à l'objet durant une certaine période : il y a donc distinction entre les durées de vie de l'objet et de l'essai, ceci permet **la flexibilité du système envers les données et leur lien géographique**, toute donnée étant toujours référencée directement ou par liaison à l'aide des identifiants, l'identifiant automatiquement généré de ces tables étant appelé « code » et non ID,

- Les **tables ESSAIS** : elles n'ont rien de spécial si ce n'est la génération automatique d'un ID daté permettant d'identifier chaque ligne de la table, permettant, là aussi, **une souplesse du système**.

9.4 Digitalisation et correction des données spatiales

Pour cette partie consulter, l'annexe 9 qui illustre quelques résultats de la digitalisation par l'image.

9.4.1 La digitalisation par rapport à l'orthophoto et le cadastre

La digitalisation par rapport à l'orthophoto permet de digitaliser les parcelles d'essais ainsi que toutes les infrastructures du domaine. **Cette digitalisation est référencée au territoire**, donc fidèle de la réalité, à quelques mètres près. On estime la précision entre la digitalisation et la position cadastrale des objets à environs 3 mètres. Cette précision est suffisante pour le projet, une autre acquisition, tel qu'un relevé topographique, ne serait pas justifiée économiquement et apporterai une précision non nécessaire. Le cadastre sert en outre de calage du périmètre extérieur du domaine de Caudoz.

9.4.2 La digitalisation semi-topologique

L'annexe 11 décrit la méthode pas à pas afin de réaliser une digitalisation rapide et conforme des éléments du système d'expérimentation.

La digitalisation semi-topologique concerne la digitalisation des éléments du système d'expérimentation. En effet, leur digitalisation s'opère à partir d'un plan papier réalisé « à la règle » ou sur informatique grâce à EXCEL. Il n'y a donc pas de possibilité de référencer ces objets : la seule possibilité serait de réaliser un relevé topographique, encore une fois non justifié pour les mêmes raisons que précédemment. En effet, la position des lignes de vigne importe peu du moment que celles-ci sont localisées de manière semi-topologique. On entend par-là que les éléments du système d'expérimentation soient localisés dans leur parcelle respective, à leur place respective topologiquement sans donc être référencés de manière absolue. Une structure topologique pure reviendrait à une représentation schématique des éléments sans même des coordonnées, mais de façon plutôt ordinale que cardinale, ce qui n'est pas notre cas ici, le choix ayant été fait § 7.4.

9.4.3 La correction des données spatiales

La correction des données spatiales a lieu lorsque des tables vectorielles sont importées dans le SIG, sans digitalisation des éléments. C'est **le cas pour les unités de sol pédologique**. La correction consiste à rendre les couches importées **compatibles** avec les couches digitalisées. Cette compatibilité consiste à **rendre conforme les limites des couches** en présence. En effet, la couche « unité de sol » importée a été digitalisée sur la carte nationale CN25, alors que les couches digitalisées prennent leurs marques sur l'orthophoto et le cadastre. Il en ressort des différences au niveau des limites que l'on doit corriger. Dans cette étape, il ne faut que « ramener » l'objet de la couche « unité de sol » sur les objets de la couche « parcelle » qui est la référence, car calée sur le cadastre. Il ne faut en aucun cas modifier les limites internes des zones car il y aurait, à ce moment là, modification géographique de la donnée spatiale, ce qui n'est pas le cas pour les limites extérieures qui sont simplement rendues conformes. On peut parler d'un **« rattrapage »**

de digitalisation, réalisé avec succès grâce au SIG. Cette correction doit avoir lieu afin d'exploiter au mieux toutes les couches entre elles, notamment pour les analyses spatiales oú une **cohérence entre couches** est nécessaire.

9.5 Représentation spatiale des données dans MANIFOLD

9.5.1 Les objectifs de la représentation

Principalement, trois objectifs ont été suivis ici, la table 15 les décrit au § 9.5.2. que l'on retrouve aussi en annexe 10.

- Discerner rapidement les objets des différentes couches.
- Faciliter la visualisation et la sélection des éléments.
- Créer la convivialité de l'espace de travail.

9.5.2 Les représentations des objets au sein du SIG

COUCHE	TYPE D'ENTITE	REPRESENTATION
Zones	Surface	Neutre et contour rose
Routes	Surface	Gris à points et contour noir
Bâtiments	Surface	Hachures marron clair et contour noir
Station météo	Surface	Hachures épaisses marron clair et contour noir
Domaine	Surface	Neutre et contour jaune
Unité de sol	Surface	Beige et contour noir
Sondages	Point	Point orange
Profils	Point	Cibles rouges
Pieds	Point	Marron clair
Lignes	Ligne	Noire épaisse
Répétitions	Surface	Violet pâle et contour noir
Variantes	Surface	Couleurs vives et discriminantes au contour noir
Parcelles	Surface	Couleurs ternes et discriminantes au contour noir

Table 15 : la représentation des objets des couches au sein du SIG.

9.6 Les relations entre tables

9.6.1 Les relations selon le MLD

Les relations entre tables sont celles données par le **modèle logique des données (MLD)** (Cf. annexe 6 : le MLD). Les identifiants sont les moteurs des liaisons entre tables.

9.6.2 Les relations logiques pour les chercheurs

Afin de se « retrouver » rapidement et de ne pas perdre de vue les identifiants logiques habituels, les chercheurs ont souhaité avoir dans chaque table les éléments [Domaine,numéro parcelle, titre essai] obtenus par liaison à la table « base parcelle » **(jointure**) . D'autres relations sont créées au besoin pour **clarifier la lecture par jointures entre tables** (Cf.§ 10.1.3.3 : partie sur les jointures).

9.7 Les scripts réalisés et leurs rôles

9.7.1 Qu'est-ce qu'un script ?

Un script est, dans MANIFOLD et généralement en informatique, un **petit programme réalisant une fonction donnée**. Le script est écrit dans le langage de programmation du logiciel. On l'exécute par un *recompute* dans MANIFOLD, ce qui en fait une fonction « à la demande de l'utilisateur ». Les scripts permettent de réaliser des choses que le logiciel ne fait pas de manière programmée par défaut. Dans le présent projet, on utilise quatre scripts, très courts, qui aident pour beaucoup dans la conception du prototype, notamment au niveau de la création des identifiants.

9.7.2 Les scripts de concaténation

Une concaténation consiste au **regroupement de plusieurs champs en un seul**, c'est en quelque sorte une agglomération de plusieurs mots à la suite, séparé ou non par des «- , /… ».

Dans le projet, ces scripts servent à la table « météo » et pour toutes les tables « essais ». Il s'agit de concaténer les ID des tables de base avec l'année d'essais pour obtenir un autre ID du type : txxrxx-xxxx.(txxrxx : ID de la table de base correspondante – xxx : année d'essais). (Cf.§8.5).

```
Function concatenation
        concatenation = (record.data("Code_variante"))&("")&(record.data("Année"))
End Function
```

9.7.3 Les scripts de calcul de surface

Le calcul de la surface se réalise par un **script qui limite la précision au mètre**. Il s'exerce sur les parcelles, les unités de sol, le domaine…afin de donner des indications sur la surface des éléments. En revanche, il ne s'applique pas sur les éléments du système d'expérimentation qui ne sont digitalisés que de manière semi-topologique ! Le calcul des surfaces se fait dans ce cas, par les chercheurs, par mesure sur le terrain lors de la création de l'essai.

```
Function surfacemetrecarre
        surfacemetrecarre = round(record.data("area"))
End Function
Function surfacehectare
        surfacehectare = surfacemetrecarre/10000
End Function
```

9.7.4 Les scripts de création d'ID

La **création d'ID unique dans toute la base** se réalise en prenant le numéro de la table incluse dans le logiciel précédé de « t » en la concaténant avec le numéro de la ligne (record) précédé de « r ». On obtient donc un ID du type : txxryy, avec xx le numéro de la table et yy celui de la ligne de cette table. (Cf.§8.5). Ceci se réalise sur toutes les tables BASE.

```
Function iduniquebasevariante
        iduniquebasevariante = "t"&table.id&"r"&record.id
End Function
```

9.7.5 Les scripts d'affichage des coordonnées

Ce script réalise **l'affichage des coordonnées (Y ;X), en les limitant au mètre**, des données ponctuelles que sont les profils et les sondages, afin de pouvoir les retrouver sur le terrain.

```
Function Xcoordinate
        Xcoordinate = round(record.data("X (I)"))
```




```
End Function
Function Ycoordinate
        Ycoordinate = round(record.data("Y (I)"))
End Function
```

Résumé des points importants :

Le logiciel SIG choisi est MANIFOLD, le système de référence CH1903 oú sont importées et digitalisées les différentes couches. Le projet s'organise en dossiers qui regroupent les couches selon le MCD. La digitalisation s'appuie sur l'orthophoto calée grâce au cadastre. Au contraire, la digitalisation des éléments du système d'expérimentation est semi-topologique. Les relations entre tables permettent d'effectuer les relations du MCD et des scripts permettent d'effectuer des fonctions sur les attributs des tables.

Dans ce chapitre :
Nous montrons comment le SIG parvient à réaliser les besoins et attentes des chercheurs.
Nous mettons en évidence ceci par quelques exemples et illustrations.
Enfin, des scenarii d'utilisation du SIG envers les données d'essais sont créés, fictivement, afin de montrer la souplesse, la flexibilité et l'adaptation du SIG à cette situation particulière. .

10 Analyse, exemples et scenarii

10.1 Exemples et scénarii

10.1.1 Visualisation géographique

10.1.1.1 Des éléments du domaine

L'annexe 9.partie.2 illustre les éléments du domaine. Le § 9..5 donne les explications de leur représentation.

10.1.1.2 Des parcelles d'essais

L'annexe 9.partie.1 illustre les éléments du domaine. Les parcelles apparaissent dans des couleurs différentes sans « overlap » afin de pouvoir les distinguer. Le ton des couleurs est terne.

10.1.1.3 Des éléments du système d'expérimentation

Les éléments du système d'expérimentation ont été mis en évidence au § 9.4 et §9.5 ainsi que dans l'annexe 9, oú l'on visualise la structure de ces éléments formant une partition du territoire. On rappelle que ces éléments sont digitalisés de façon à ce que aucun élément n'en cache un autre !

L'annexe 12 présente la parcelle 9, l'annexe 13 la parcelle 17.

10.1.1.4 Création de labels

Les labels sont des attributs stockés dans des tables qui apparaissent sur leurs éléments géographiques respectifs. Ceci permet d'avoir une étiquette sur les objets dont on désire connaître le nom ou un renseignement. L'annexe 14 illustre ces labels. Ils ont été créés sur tous les éléments du système d'expérimentation et permettent ainsi aux chercheurs de se retrouver. Les couleurs ont aussi ce rôle : chaque élément n'a pas son label, surtout si l'objet est dissocié, comme les variantes par exemple. Ceci évite une surcharge graphique.

10.1.2 Visualisation et sélection au niveau des tables

NOTE : LES DONNEES ONT ETE INTRODUITES DE MANIERE LA PLUS FIDELE POSSIBLE POUR L'ANNEE 2001. TOUTEFOIS, CERTAINES D'ENTRES-ELLES *S'INSPIRENT* DE LA REALITE AFIN DE POUVOIR ILLUSTRER LES DONNEES MANQUANTES POUR CAUSE DE PIQURE ACETIQUE.

10.1.2.1 Visualisation des données tabulaire

Afin d'illustrer le passage des données brutes au données SIG, un exemple est donné en annexe 15. Elle illustre les données brutes de vendange issues du cahier de vendange de VITIDATA pour un essai de greffage herbacé. Le choix de saisir les données par ligne ou directement par variante est démontré.

10.1.2.2 Visualisation des données géographiques et de leur correspondant tabulaire

L'annexe 16 montre comment faire rapidement les liaisons entre les tables OBJET / BASE /et ESSAIS afin de consulter et visualiser en même temps les données, même si plusieurs tables différentes sont nécessaires.

10.1.3 Requêtes

10.1.3.1 A l'aide du menu barre *select*

L'annexe 16 illustre ce menu barre situé dans la barre des tâches du logiciel MANIFOLD. Des sélections simples sont possibles, aussi bien pour des objets géographiques que pour des attributs de tables.

10.1.3.2 A l'aide de scripts de requêtes

Les scripts sont définis § 9.7.1. Ici, ils jouent le rôle de recherche d'éléments. Ils sont appelés *query*. L'utilisateur peut ouvrir le *query* et changer le choix de sa sélection (il remplace x, y et z par les numéros qu'il

recherche).(Cf. fig. 15) Ces requêtes permettent des sélections plus fines. Six scripts ont été réalisés pour répondre aux questions les plus fréquentes des chercheurs. (Cf. table 16)

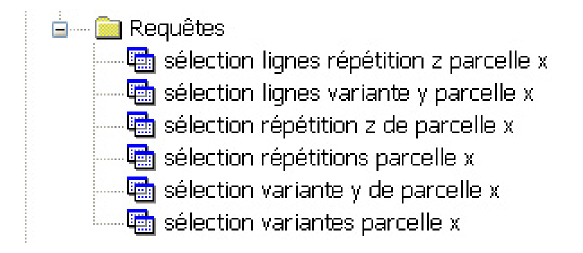


Figure 15 : les six requêtes prédéfinies grâce à un script.

```
REQUETE: SELECTION VARIANTES PARCELLE X

select *
from [Base variante]
where ([Numéro_parcelle]=20)

REQUETE: SELECTION REPETITIONS PARCELLE X

select *
from [Base_repetition]
where ([Numéro_parcelle]=17)

REQUETE: SELECTION VARIANTE Y DE PARCELLE X

select *
from [Base variante]
where ([Numéro_parcelle]=20)and([Numéro_variante]=Aa)


REQUETE:SELECTION REPETITION Z DE PARCELLE X

select *
from [Base_repetition]
where ([Numéro_parcelle]=20)and([Numéro_répétition]=1)

REQUETE:SELECTION LIGNES VARIANTE Y PARCELLE X

select *
from [Base ligne]
where ([Numéro_parcelle]=20)and([Numéro_variante]=Aa)

REQUETE:SELECTION LIGNES REPETITION Z PARCELLE X

select *
from [Base ligne]
where ([Numéro_parcelle]=20)and([Numéro_variante]=Aa)
```

Table 16 : les six scripts réalisés à des fins de requêtes selon choix.

10.1.3.3 Grâce à des jointures entre tables

Une jointure correspond au regroupement des attributs d'une table dans l'autre grâce à une relation entre deux attributs semblables. On peut ainsi par exemple joindre à la table « résultats principaux de vendange par variante » (Cf. Fig. 16) les attributs [Dégrappage, Poids_grappes, Poids_baies, Poids_bois_taille, Pourriture_raisin, Fertilité] de la table « résultats principaux données de contrôle par variante » (Cf. Fig. 17) par le biais de leur numéro d'essai [Numéro_essai]. (Cf. Fig. 18).

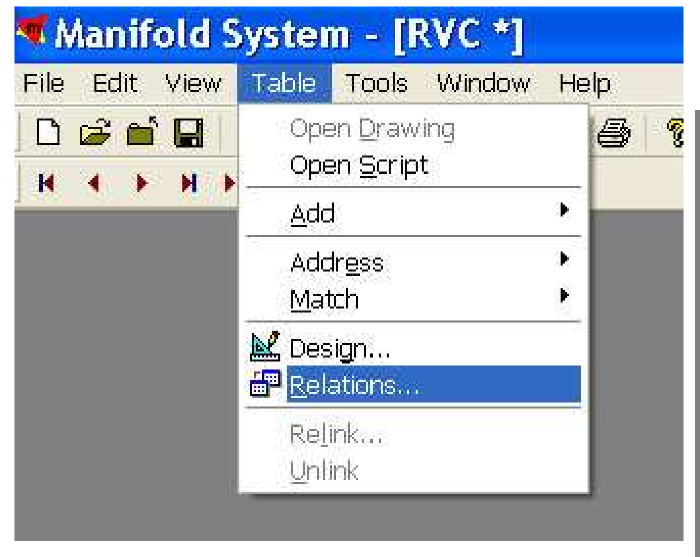


Figure 16 : comment créer une relation entre deux tables ?

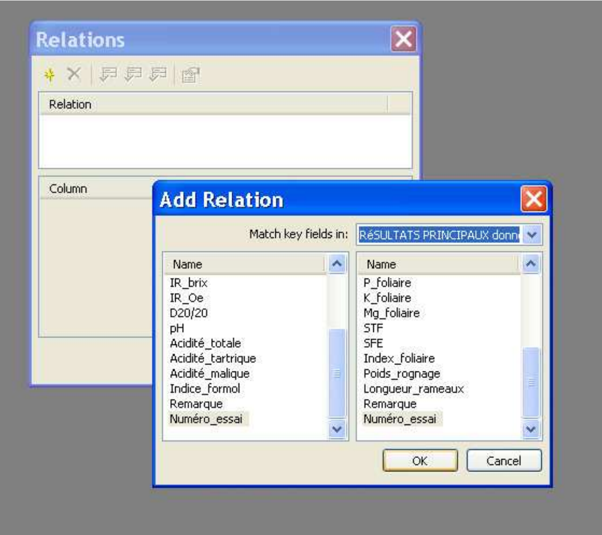


Figure 17 : opération de jointure d'attributs d'une table à l'autre.

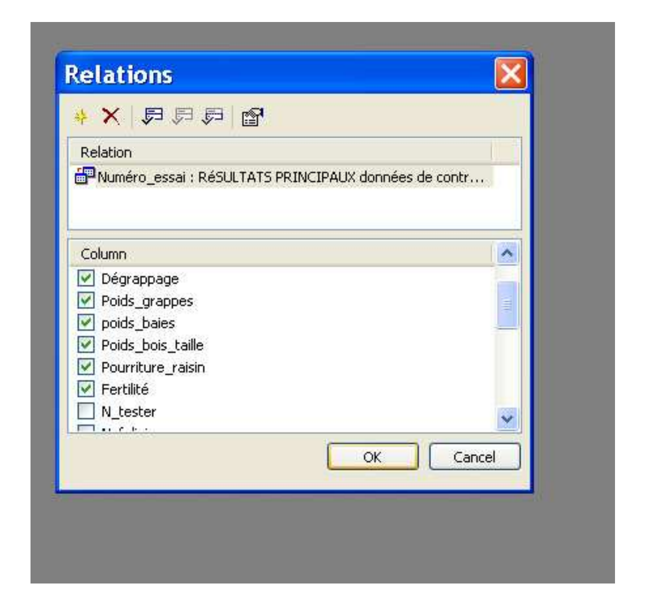


Figure 18 : choix de l'attribut de liaison lors d'une jointure de table.

10.1.4 Cartes thématiques

Une **carte thématique** est une carte réunissant des données par sujets identiques ou par types de résultats ce qui donne le thème de la carte. Ici cinq cartes thématiques ont été réalisées, annexe 17, ce sont :

- Carte thématique des unités de sol pédologique,
- Carte thématique des profondeurs de sol,
- Carte thématique de la réserve hydrique modale (RHM) des sols,
- Carte thématique de la réserve utile des sols (RU),
- Enfin la carte thématique effectuant le croisement entre le type de roche mère et le type de sol.

Ces cartes thématiques sont munies :

- D'une légende,
- D'une orientation,

- Et d'une échelle.

Elles illustrent les thématiques les plus fréquemment utilisées par les chercheurs. Le choix d'une **sémiologie graphique** a été réalisé pour chaque carte. Cette sémiologie se veut simple et intuitive (jeux de couleur marron pour le sol, et bleu pour les teneurs en eau, allant des couleurs foncées aux couleurs claires lors de la décroissance du phénomène représenté).

10.1.5 Edition de documents

10.1.5.1 Par impression directe

- Pour des données graphiques : il est possible d'imprimer ce qui est vu à l'écran par une impression directe. Il est ainsi possible de créer toute sorte de vues sur des éléments. Il est toutefois préférable d'imprimer depuis une *map* : on peut choisir les couches visibles, leur ordre de succession, …, mais aussi d'introduire les labels créés, le menu *view* permettant l'introduction de légendes, échelles, orientations…. Un exemple est disponible annexe 18.

- Pour les données tabulaires : il est possible d'imprimer directement une table, le format d'impression est convivial et permet ainsi une lecture aisée des données. (couleurs alternées pour différencier les lignes de données). Un exemple est disponible annexe 19

10.1.5.2 Par *layout*

Le *layout* est un masque d'impression, prédéfini, des données graphiques. Il permet des impressions avec titre, échelle, légende, orientation,…, inséré automatiquement.

- Pour des données graphiques : Dans le logiciel MANIFOLD existent deux types de *layout* : les *layout* « fixes » et les *layout* « libre ». Le l*ayout* fixe imprime une carte définie : c'est la carte du domaine pour notre exemple annexe 20.3. Le *layout* libre permet d'imprimer ce que l'on choisi par sélection : ce sont des zooms parcellaires ayant pour titre libre « Pully, Parcelle…, Répétition…, Variante… » que le chercheur complète selon son choix. Voir à ce sujet l'annexe 20.1 et 20.2.

- Pour les données tabulaires : il est définit par défaut, voir § précédent 10.2.5.1.

10.1.5.3 Par exportation

- Pour des données graphiques : on peut toujours exporter des données graphiques dans un autre format par le menu d'exportation de MANIFOLD. Le format proposé est le format d'échange de données reconnu .dxf mais sert surtout pour les exportations / importations entre SIG. Une capture d'écran permet d'obtenir une image (.tif, jpg…), ce qui peut rendre service pour introduire une image dans un document tel un rapport.

- Pour les données tabulaires : l'annexe 25 donne l'exemple de l'exportation d'une table en format EXCEL, utilisé par les chercheurs pour le traitement des données. Des importations de tables sont aussi possible depuis plusieurs formats, mais attention à garder une rigueur tant au niveau de la hiérarchisation du projet que des attributs contenus dans la table !

10.1.6 Outils SIG : l'exemple du « containing »

Un outil SIG est un outil implanté par défaut dans le logiciel qui permet **d'effectuer une tâche définie**. L'exemple choisi ici est l'outil *containing*. Il permet de **sélectionner des objets contenus dans d'autres**.

Ceci est utile pour les éléments du système d'expérimentation qui forment une partition du territoire oú les objets s'englobent les uns les autres. (La parcelle contient les répétitions et les variantes, les répétitions et les variantes contiennent les lignes, les lignes contiennent les pieds). On imagine facilement les applications qui en découlent : un exemple est donné annexe 21.

10.1.7 Souplesse du SIG au niveau des données d'essais

L'annexe 26 illustre cette partie, grâce à la parcelle 111 fictive créée pour démonstration. Nous ne pouvons pas illustrer tous les cas de figures recensés, mais nous définissons ci-dessous trois grandes catégories auxquelles appartiennent ces cas.

Cette souplesse est due, rappelons-le, grâce à la distinction entre les OBJETS de leurs données de BASES, et de leurs données d'ESSAIS : Cf. § 8.3. Cette distinction a eu lieu lors de la conception du modèle. La modélisation a pris en compte les besoins et attentes des chercheurs en ce sens, en voici le résultat.

10.1.7.1 Possibilité d'adaptation des essais en cours d'expérimentation

Par adaptation des essais en cours d'expérimentation, il faut comprendre que le SIG offre la possibilité aux chercheurs de modifier un essai, au niveau de ses données de base, comme il le souhaite. Les données ESSAIS sont aussi modifiées par l'ajout d'une ligne à la table renvoyant à l'essai modifié. L'historique, c'est à dire les informations concernant l'essai précédent, est conservée. Ceci est valable quel que soit l'OBJET du système d'expérimentation.

10.1.7.2 Mobilité spatio-temporelle des objets

La mobilité spatio–temporelle des objets contient la mobilité dans l'espace et aussi dans le temps. Ceci a impliqué la création d'un attribut, sur les tables OBJET, nommé *actif*, permettant de savoir si l'OBJET est en cours d'essai actuellement ou non. (Cf. § 8.3) . Si l'OBJET change, alors on a affaire à un nouvel objet. Ce nouvel objet comporte un nouvel ID, et des attributs qui lui sont propres. Par relation, les attributs des tables liées à l'OBJET, indiquent le nouvel objet et non l'ancien. Signalons pour terminer, que l'ancien objet existe encore, il n'est tout simplement plus actif.

Voici trois cas de figure, les plus courants :

- **La création d'objets** : on crée un nouvel élément du système d'expérimentation,
- **La fusion d'objets** : on fusionne deux parcelles en une seule,
- **Le déplacement d'objets** :on déplace la limite d'une parcelle, au profit de l'une et au dépend de l'autre.

10.1.7.3 Autres volontés particulières

D'autres volontés particulières, nouvelles, rares…, peuvent être comblées en s'appuyant sur les relations entre tables OBJETS / BASES / ESSAIS. Une recherche peut alors avoir lieu afin de retrouver tout ce qui concerne un élément, ce qui simplifie les choses. A ce stade, le modèle mis en place au sein du SIG, doit encore être validé, amélioré, afin de savoir si ce dernier rempli bien toutes les tâches, même celles auxquelles nous n'avons pas pensé.

10.1.8 Analyse spatiale

10.1.8.1 Visuelle

Une analyse spatiale visuelle est l'analyse de deux paramètres croisés graphiquement grâce à une sémiologie graphique judicieuse (par exemple des couleurs différentes mélangées avec une densité de points).

Ici, deux cartes permettant une analyse spatiale visuelle ont été créées. Il s'agit de : (Cf. annexe 22).

- La réserve hydrique modale (RHM) croisée avec le poids des bois de taille : visuellement, donner une interprétation d'une **éventuelle relation entre les deux paramètres** est difficile. Dans le § 10.1.8.2 suivant, nous tâcherons d'analyser cette éventuelle relation par le calcul d'un coefficient de corrélation.

- La réserve hydrique modale (RHM) croisée avec le poids des baies : aussi ici, donner une interprétation est difficile, ceci pour l'entier du domaine. Toutefois, un zoom sur la parcelle 27 permet de mettre en évidence que les baies sont plus lourdes là oú la RHM est plus élevée. Cependant, cette observation n'est que locale est ne peut donc pas être validée. Un calcul s'impose donc pour justifier cette observation, que nous pouvons qualifier **d'hypothèse** à ce stade.

10.1.8.2 Par calcul d'un coefficient de corrélation

Dans cette partie, l'analyse entre la réserve hydrique modale (RHM) et le poids des bois de taille est **réalisée mathématiquement par calcul d'un coefficient de corrélation**. L'annexe 23 met en évidence les résultats graphiques de la méthode, alors que l'annexe 24 met en évidence les calculs et le résultat obtenu.

L'objectif est de rechercher une éventuelle relation entre deux paramètres définis dans l'espace. Ceci est appelé **corrélation**. Ici, nous testons les potentialités offertes par le SIG en terme de traitement, la corrélation que nous recherchons grâce au SIG existe et est démontrée.
Les paramètres étudiés sont :

1 : **la Réserve Hydrique Modale (RHM)** : elle représente la teneur en eau potentielle du sol, modale signifiant « valeur la plus probable dans la zone ». Ce paramètre est disponible sur l'entier du domaine du Caudoz. L'unité est le millimètre [mm]. Chaque « unité de sol » comporte une valeur de RHM.

2 **: le poids des bois de taille** (les sarments) : il représente le poids des bois de taille de la vigne, on appelle ceci les sarments. L'unité est le gramme par sarment, [g]. Ce paramètre est défini de manière zonale, en fonction des données disponibles. Ainsi, une valeur moyenne peut être fournie pour une parcelle entière ou bien simplement pour une répétition, ce qui est le cas le plus fréquent. Voir les données en annexe 24…

La situation prise en compte comporte plusieurs facteurs qui tendent à **amoindrir la relation entre les paramètres**, ce sont :

- Situation non extrême : la station de Pully ne fait pas partie des cas de figure extrêmes, ou la relation apparaît comme évidente, comme par exemple dans les vignobles du Lavaux ou les précipitation ne sont que de 600 à 800 mm / an (Pully : 1100mm/an, et année 2003 : 800mm),
- Choix de l'année 1997 : 1997 est une année bien arrosée, qui conduit ainsi à un lissage des valeurs de RHM sur le domaine,
- Particularité des terres de Pully : ces terres constituent des bonnes réserves en eau, qui conduisent à des poids de grappe et des poids de bois de taille importants,

- Petit domaine exploratoire : la surface utilisée pour étudier la corrélation ne constitue pas une surface exploratoire assez importante, elle n'est en plus que légèrement variée,
- Echantillonnage faible : l'échantillonnage des unités de RHM est précis à une dizaine de mètres environ, et les valeurs des poids de bois de taille ont été relevées par répétition, ce qui constitue des moyennes par zones.

Nous sommes ainsi dans une situation non évidente, où le calcul d'un coefficient de corrélation apportera un aspect mathématique.

PRECAUTION À PRENDRE : les deux paramètres doivent être **cohérents ensemble**. C'est-à-dire que les paramètres doivent pouvoir être analysés ensemble dans **des conditions homogènes et identiques sur toute la surface d'étude commune**. Dans notre cas, nous devons être attentifs à cette précaution. En effet, le système d'expérimentation conduit à la réalisation d'essais différents sur chaque parcelle oú les mêmes paramètres (attributs) sont relevés pour chaque essai. Il est donc faux de rechercher une corrélation entre paramètres homologues mais provenant de conditions d'expérimentation différentes.

METHODE SUIVIE : Cette relation peut être déterminée en **croisant les deux couches** contenant chacune les valeurs (et non les classes) de leur paramètre respectif. Ce croisement s'effectue par ***polygone overlay***. Ceci revient à trouver le plus petit polygone commun à chaque couche, et de lui attribuer la valeur correspondante pour le premier paramètre et de même pour le second. Ensuite, le calcul d'un **coefficient de corrélation,** ϱ, entre les couples de valeurs (1 ;2) ainsi formés, donne une indication sur la part de relation entre les deux paramètres. Si ce coefficient vaut 1, ceci signifie que les deux paramètres sont liés par une relation, ils sont donc dépendants. Au contraire, s'il vaut 0, il n'y a aucune relation entre eux. La valeur (-1) signifie une dépendance totale opposée.

ϱ=(COVARIANCE(1 ;2)/(VARIANCE(1)*VARIANCE(2))^(1/2))

Soit :

$$\varrho=(\sigma(1;2)/\sigma(1)*\sigma(2)) \quad \text{dans } [-1;1]$$

RESULTAT : Un coefficient de **0.45** est obtenu en analysant les couples de valeur dans leur totalité. En enlevant 6 couples de valeurs marginales sur 41 couples, représentant ainsi 16% de la surface d'analyse, nous obtenons un coefficient de **0.77**.

INTERPRETATION : les facteurs tendant à amoindrir la relation ne jouent pas un rôle déterminant dans le calcul. Ce sont plutôt les **valeurs marginales** (et non aberrantes, car elles existent !) qui conduisent à deux **interprétations différentes**, que nous les prenions en compte ou non.

CONCLUSION : l'**échantillonnage** est la cause principale de cet écart entre coefficient de corrélation. Un échantillonnage, par surface de variante individuelle doit être pris en compte afin d'obtenir une population plus représentative d'un phénomène très local.
Malgré tout, le SIG montre sa réelle potentialité : si l'échantillonnage est en adéquation avec ce qu'il représente, alors la corrélation est véritablement illustrée et démontrée. Les perspectives sont nombreuses, notamment dans les situations comme celle-ci où l'échantillonnage ponctuel coûte cher et prend beaucoup de temps.

10.2 Analyse de fonctionnement, apports et limites

En se reportant à la figure 9 du § 5, qui traçait le bilan de l'évaluation des besoins des chercheurs, on peut dire que le SIG a accomplit sa mission. De nouveaux besoins risquent de voir le jour. En effet, l'apport d'une nouvelle technologie crée des besoins, on dit communément : *l'offre crée la demande*. Ceci est vrai pour l'édition de documents, la réalisation des cartes thématiques… Pour ce qui est le la base de données elle-même, elle doit faire ses preuves au sein de l'entreprise, dans des conditions d'exploitation réelles.

Le SIG n'apporte cependant pas une réponse à tous les détails qui existent, et qui font de chaque entreprise une individualité. Le remplacement des bases de données existantes n'est pas envisagé, (Cf. § 6 perspectives de développement) mais le SIG ne pourra pas, par exemple, remplir toutes les fonctions, apportées par les différentes bases de données, en un seul produit.. Un système automatisé, comme l'est VITIDATA, ne peut l'être pour l'ensemble des bases de données (édition de codes barres, maintenance, édition de cahiers…). Le SIG prend donc **sa force** dans :

- Son **pouvoir de centralisation de l'information**, pas réalisée jusqu'ici,
- Ses possibilités **d'édition de documents géographiques**,
- Sa **gestion de l'historique**,
- Sa **gestion des éléments connexes**,
- Ses apports en nouveautés (Cf. § 8.6), dont les **possibilités de consultation interannuels et intra- / inter domaine**, nouveauté résultant de la centralisation et de l'accès aux données, permettant, soyons certain, d'augmenter l'apport d'informations ciblées sur des essais de même type.

Ses **limites** ne sont que temporaires, des solutions doivent être trouvées ultérieurement, pourquoi pas lors d'un prochain travail de diplôme ? (Cf. 12.2) :

- La **difficulté des liaisons entre tables** qui devraient être automatisées au sein d'une même thématique fonctionnelle, à définir selon les besoins, pour permettre aux chercheurs de consulter toutes les données ayant à faire avec le même identifiant,
- **L'absence de masques de saisie**, qui devraient permettre la convivialité du système, encore plus qu'il ne l'ait actuellement,
- La **difficulté de réaliser une analyse spatiale** avec ce logiciel, qui devrait être programmer, si cela est possible, afin de rendre la manipulation assistée.

Cela dit, les implications extérieures de ce SIG sont importantes et multiples. En effet, **le SIG apporte la possibilité de gérer un domaine**. Cette gestion est assez libre, elle est simplement guidée. Le domaine en question, peut être un domaine de recherche, un domaine viticole, une grande cave (traitant avec plusieurs domaines), une coopératives de vignerons… **Le modèle mis en place doit être soit adapté à un cas précis, ou bien, dans la plupart des cas, utilisé selon les besoins de l'entreprise**. Ainsi, un vigneron n'utilisera que la partie *parcelle* des données d'essais, il pourra se créer, tout de même, quelques répétitions et variantes pour ces propres recherches ou pour créer sa propre base de données afin de mieux comprendre, mieux connaître, mieux exploiter son domaine. Toute la filière est alors touchée.

En plus de cette considération sur les implications extérieures dues au SIG, il faut rappeler que le SIG a, tout d'abord, été créé pour **être extensible aux autres domaines viticoles de recherche, mais aussi aux autres domaines de recherche ayant une relation référencée sous la forme du système d'expérimentation**.



Dans le même registre, il ne faut pas oublier que le SIG apporte, ou peut être joint, à des **bases de données complémentaires**, externes, apportant leurs petits plus, qui deviennent vite nécessaires : la centralisation de l'information ne peut faire que du bien si elle est gérer avec discernement.

Résumé des points importants :

Le SIG permet d'exécuter beaucoup de tâches communes, ordinaires mais nouvelles : visualisation, consultation, recherche, interrogation, édition, impression, fonctions…

Le fonctionnement du SIG répond aux besoins et attentes des chercheurs, il permet de **gérer divers domaines** de diverses natures par adaptation de sa configuration ou tout simplement par utilisation ciblée des éléments le composant. **On démontre donc ici la possibilité, à partir d'un même modèle, de gérer des cas similaires.**

Dans ce chapitre :
La traçabilité en générale aujourd'hui.
La traçabilité pour les données d'essais et le vin produit grâce au SIG, son processus, sa structure et son contenu.

11 Notion apparentée : la traçabilité

Il est intéressant, pour terminer ce projet de travail pratique de diplôme, d'aborder la notion de traçabilité. En effet, ce travail a démontré le potentiel certain d'un SIG dans une organisation de recherche et il est maintenant temps de s'occuper de la notion de traçabilité, de voir comment celle-ci s'intègre grâce au SIG.

11.1 Qu'est ce que la traçabilité ?

Traçabilité est un mot nouveau. Nous entendons par traçabilité l'aptitude à retracer l'historique, l'utilisation ou la localisation d'une entité au moyen d'identifications enregistrées. La traçabilité permet de suivre et donc de retrouver un produit ou un service depuis sa création (production) jusqu'à sa destruction

(consommation). La traçabilité n'est pas définie par une loi mais par une norme, à savoir la norme internationale ISO 8402. Les prochaines normes laissent d'ailleurs entrevoir un futur « tout tracé ».

Ce mot est à la mode actuellement. En effet, l'incertitude alimentaire, le besoin de savoir d'où vient le produit, le besoin de poursuivre et d'attribuer les responsabilités en cas de dommage et la poussée non négligeable du BIO dans les sociétés modernes, poussent les consommateurs à toujours en savoir d'avantage. Ceci se comprend aisément par les derniers scandales : poulet à la dioxine, vache folle, poisson d'élevage au prix fort, blé et mais transgénique vendu pour du biologique, et enfin pour ce qui nous concerne plus, des millésimes mélangés à des vins de table !

Pour notre projet, la traçabilité prend une autre tournure, moins alarmante, mais se devant d'être présente pour assurer le suivi des grappes à la bouteille destinée à la dégustation. Il en va ainsi de la survie des essais et de leur réussite, notamment pour les bilans interannuels. Il faut aux chercheurs de la RVC garantir la traçabilité de leurs essais, pas simplement de la production à la consommation, mais à chaque étape de la saisie de données ainsi qu'après ces étapes, afin de conserver un historique, qui leur sera utile pour établir des conclusions d'essais. (Cf. § 4 : le fonctionnement de la RVC)

11.2 Qu'est-ce que permet la traçabilité ?

L'objectif premier de la traçabilité est de pouvoir identifier un produit, un lot de produits ou encore un service afin de pouvoir le retirer très rapidement et avec un maximum de sécurité en cas de non-conformité, de danger. La traçabilité offre également l'avantage de pouvoir intervenir en amont de la distribution en permettant par exemple de contrôler la qualité du produit depuis l'origine des ses matières premières. Ce qui autorise une nette diminution des coûts de non qualité intervenant traditionnellement sur les produits finis .De nombreuses statistiques peuvent ressortir d'une traçabilité, très utile pour un service après vente ou un service marketing. Les flux de matières premières, de produits finis sont également mieux identifiés. En résumé, la traçabilité permet d'améliorer la qualité, le service et l'efficacité globale d'une entreprise. [Tiré du site www.tracehabil.com : société française spécialisée dans l'identification et la traçabilité].

Ceci correspond tout à fait à la RVC qui a :

- Besoin d'identifier, (les différentes thématiques des données d'essais)
- Et de conserver les données d'essais, (historique)
- Dans le temps, (bilans interannuels)
- Et dans l'espace, (les bouteilles de dégustation).

Pour notre cas, la traçabilité doit permettre de faire le lien entre les bouteilles destinées à la dégustation et l'ensemble des autres données thématiques dont le tout compose la réalisation d'un essai.
A l'heure actuelle, les données sont réunies par essai et par variante dans leurs cahiers respectifs (Cf. § 4 : le fonctionnement de la RVC) avec leurs identifiants respectifs différents selon le thème, permettant de **faire ainsi le lien entre les bases de données** :

- Le numéro et / ou le titre de l'essai,
- Le code et / ou le titre de la variante,
- Le lieu de production,
- L'année d'essai,

Les bouteilles de dégustation sont stockées en cave. Elles comportent les indications suivantes :
- L'année de production,
- Le lieu de production,

- Le titre d'essai et la variante considérée,
- Le numéro d'essai figurant dans le cahier de dégustation, (exemple : 98-115, permettant d'identifier la variante par le biais de la table des matières du cahier).

On reporte donc sur la bouteille les identifiants logiques permettant de connaître les **données antérieures,** c'est-à-dire celles qui ont donné le vin, et les **données postérieures**, c'est-à-dire les données de dégustation.

11.3 Comment la traçabilité est –elle assurée grâce au SIG ?

Il est nécessaire de rappeler à cette étape, que le SIG forme un ensemble : **processus, structure et contenu**. Ainsi, le SIG peut assurer la traçabilité grâce à des outils, des fonctions (processus), grâce à la forme de conception de la base de données qu'il abrite (structure) et enfin grâce à l'ensemble des données dont il dispose (contenu).

11.3.1 Traçabilité des données d'essais

La traçabilité des données d'essais est assurée par le lien entre les identifiants des différentes tables thématiques des éléments du système d'expérimentation référencés. Cette référence permet de connaître la position, donc la provenance, des données, et ceci même si l'objet n'existe plus. (Conservation de l'historique ; l'attribut « *actif* » de l'OBJET vaut alors « non », mais l'OBJET existe bel et bien dans le SIG même s'il n'est pas visible à l'écran, et des données peuvent toujours s'y rattacher). Les liens entre tables permettent de passer d'un thème à un autre(Cf. §10.2.7). La traçabilité est donc assurée entre tables et géographie. L'annexe 16 illustre ces liens.

11.3.2 Traçabilité du produit : le vin

Les bouteilles de dégustation contiennent le vin, elles doivent aussi permettrent la traçabilité du produit. Ainsi, sur la bouteille contenant, rappelons-le, le vin issu de la vinification d'une variante d'un essai pour un domaine, **il faut apposer le code de cette variante** (*code variante*). En effet, cet identifiant permet de faire le lien entre les tables et donc de retrouver toutes les informations se rapportant au vin contenu dans cette bouteille. Exemple : le code t94 r5 devrait être apposé sur la bouteille de la parcelle 17, pour l'essai de greffage, variante C « gamay herbacé » , ….bien sûr, les autres identifiants logiques devraient être conservés afin de permettre aux chercheurs de garder leurs habitudes..

Résumé des points importants :

La notion de traçabilité doit être assurée que cela soit au sein de la RVC ou dans la vie courante aujourd'hui. Le SIG permet d'assurer cette traçabilité correctement par le biais de ses identifiants faisant les relations entre tables. L'introduction sur l'étiquette des bouteilles destinées à la dégustation du « *code variante* » permet alors d'assurer la traçabilité.



Dans ce chapitre :
Les conclusions de ce projet.
Les perspectives futures.
L'encouragement à poursuivre le projet.

12 Conclusion et perspectives

12.1 Perspectives futures

Pour l'implantation du SIG au sein de la RVC, les étapes suivantes sont à prévoir :

- Digitalisation complète des éléments du système d'expérimentation,
- Validation du modèle mis en place, discussion quant à des changements de vision du problème,
- Adaptation fonctionnelle du prototype à la station de recherche,
- Construction de masques de saisie afin de rendre plus conviviale le système,
- Programmation des consultations les plus courantes,
- Maintenance du système,
- Formation des collaborateurs,

Le prototype réalisé constitue un premier pas largement dégrossit, mais pas encore parfait comme pourrait le faire un professionnel de la chose. Il faut donc, à mon avis, continuer ce projet, l'améliorer, augmenter ses capacités et ses ouvertures. (Cf § suivant).
Le § 6 revient plus amplement sur les perspectives future du SIG : son mode de conception s'en inspire et le résultat est concluant. Toutefois, **la navigation au sein des tables**, notamment entre identifiants, doit être améliorée afin d'être d'une utilisation plus directe, ceci pourra, peut être, exister sous forme de masques de saisie, de création de tables pour d'autres services fonctionnant selon le même schéma conceptuel et de consultation. Ceci est **le point le plus critique du prototype réalisé**.

12.2 Encouragement à poursuivre le projet

En plus d'avoir été un projet intéressant, ce projet est porteur. En effet, il s'immerge dans la **Recherche référencée.** Son extensibilité à d'autres domaines font de ce projet un projet qui mérite d'être poursuivi afin de montrer aussi bien la généralisation de l'outil SIG que son application aux autres institutions fonctionnant selon le même schéma.

Ces institutions peuvent aussi bien être des organisations **de recherche que des vignerons, des « vignobles » ou des grandes caves**. Les extensions sont multiples et les apports peuvent avoir des implications importantes voire grandissimes.

12.3 Conclusion du projet

Après avoir suivi une démarche traditionnelle pour le monde des SIG, **les objectifs fixés sont atteints**. Le présent rapport assorti de ses annexes permet une bonne compréhension de la problématique et expose la résolution mise en œuvre. Cette résolution n'est pas unique : elle s'inspire des méthodes et des outils à disposition, elle s'est voulue **simple, extensible et modulable** afin de réaliser un SIG **souple et flexible**, **répondant aux besoins et attentes** des chercheurs, **utilisant le plus possible les données et les structures en place**. La modélisation du problème a permis de dissocier les OBJETS de leurs attributs, de permettre la mobilité spatio-temporelle des éléments du système d'expérimentation ainsi que de répondre aux volontés des chercheurs et de leurs essais.

Les deux principales difficultés de ce travail ont été, d'une part, la compréhension du fonctionnement de la station de recherche, et d'autre part, le prototypage sur un logiciel jusque là inconnu.

Nous espérons avoir fait en sorte que le SIG ait le plus de chance possible de voir le jour, sous cette forme ou sous une forme voisine (Cf. § précédents).

Pour terminer, soulignons l'importance de la collaboration entre acteurs, sans laquelle aucune solution valable n'aurait été trouvée.



Bibliographie

Ayant servie considérablement pour ce travail

DESPONDS N., JEANGROS B., TROXLER J., 1999, *Utilité d'un système d'information à référence spatiale pour un domaine agricole expérimental*, Revue suisse Agric. 31(3), 1-7.

DUBOIS Jacques, 1996, *Les vignobles vaudois : étude géographique viticole : regards sur leu passé et leur présent, prévisions quant à leur avenir*, Yens-sur-Morges, éditions Cabétia, ISBN 288295-181-7.

DUMAY Raymond,1967, *Guide du vin, le livre de poche pratique*, Paris, ISBN : 2-253-00376-X, 475 p..

GOLAY F., 1999, *SIRS 1*, notes de cours, polycopié EPFL.

GOLAY F., 2000, *SIRS 2*, notes de cours, polycopié EPFL.

MERTINA Christophe, 2003, *Avant projet de développement d'un SIG pour la gestion du domaine viticole expérimental de Caudoz*, séminaire interdisciplinaire, laboratoire de SIG, EPFL.

PLAZANET C., 2000, *SGBD*, notes de cours, polycopié EPFL.

PYTHOUD K., CALOZ R., 2001, *Etude des terroirs viticoles vaudois*, rapport d'avancement, 80p., laboratoire de SIG – EPFL, Lausanne.

RAC, 1996-1997, *Rapport d'activité*, station fédérale de recherches en production végétale, Nyon, 120p..

SPACCAPIETRA Stephano, 2000, *bases de données classiques*, notes de cours , polycopié EPFL,

ZUFFEREY V., PHYTHOUD K., LETESSIER I, FERMOND C., 2003 *Etude des terroirs viticoles vaudois : Comportement de la vigne en fonction des conditions pédo-climatiques*, rapport intermédiaire 2002, station fédérale de recherches en production végétales de Changins, Nyon, 131p..

ZUFFEREY Vivian, 2000, *Echanges gazeux des feuilles chez vitis vinifera L (CV. Chasselas) en fonction des paramètres climatiques et physiologiques et des modes de conduite de la vigne*, thèse EPFZ N°13777, Zurich, 335p..

Sites Internet

Station fédérale en production végétale : www.admin.ch/sar/rac
« De la vigne au vin » : http://users.swing.be/john/vignevin/index.htm

traçabilité: www.tracehabil.com

Pour la partie « vigne » ***bibliographie complémentaire***

CAMEFORT H. & GAMA A. *Sciences Naturelles.* Classiques Hachette. 1958.

CARMERRERE C., MADEVON D. & P. *Les vins de France : œnologie et géographie.* Nathan. 1993.

COUDERC P. *Guide pratique et historique des vins de France.* Le pré aux clercs. 1984.

ENJALBERT H. & B. *L'histoire de la vigne et du vin.* Bordas / Edition Bardi. 1987.

FOULONNEAU C. *Guide pratique de la vinification.* Armand Collin, collection verte. 1991.

FREYDIGER J. & GERVER F. *Les vins.* Edition Saud. 1972.

FRIBOURG G. & SARFATI C. *La dégustation. Connaître et comprendre le vin.* Edisud, collection de l'université du vin. 1989.

GAUTIER J.F. *Histoire du vin.* Presse universitaire de France, collection Que sais-je ? 1992.

NAVARRE J.P. & C. *Manuel d'œnologie.* Edition J.B. Baillière, collection d'enseignement agricole. 1979.

RIBEREAU-GAYON J. & P., PEYNAND E. & SUDRAND P. *Sciences et techniques du vin.* Tome 1. Edition Dumod. 1972.

Christophe MERTINA,

Lausanne, février 2004.
